Commercial Motor Vehicle Driver Fatigue, Long-Term Health, and Highway Safety: Research Needs

DETAILS

272 pages | 6 x 9 | PAPERBACK
ISBN 978-0-309-39252-5 | DOI: 10.17226/21921

AUTHORS

Panel on Research Methodologies and Statistical Approaches to Understanding Driver Fatigue Factors in Motor Carrier Safety and Driver Health; Committee on National Statistics; Board on Human-Systems Integration; Division of Behavioral and Social Sciences and Education; Transportation Research Board; National Academies of Sciences, Engineering, and Medicine

COMMERCIAL MOTOR VEHICLE DRIVER FATIGUE, LONG-TERM HEALTH, AND HIGHWAY SAFETY

RESEARCH NEEDS

Panel on Research Methodologies and Statistical Approaches to
Understanding Driver Fatigue Factors in
Motor Carrier Safety and Driver Health

Committee on National Statistics
and
Board on Human-Systems Integration
Division of Behavioral and Social Sciences and Education

and
Transportation Research Board

The National Academies of
SCIENCES · ENGINEERING · MEDICINE

THE NATIONAL ACADEMIES PRESS
Washington, DC
www.nap.edu

THE NATIONAL ACADEMIES PRESS 500 Fifth Street, NW Washington, DC 20001

This activity was supported by Award No. DTMC75-13-C-00006 from the United States Department of Transportation's Federal Motor Carrier Safety Administration. Support of the work of the Committee on National Statistics is provided by a consortium of federal agencies through a grant from the National Science Foundation (No. SES-1024012). Any opinions, findings, conclusions, or recommendations expressed in this publication do not necessarily reflect the views of any organization or agency that provided support for the project.

International Standard Book Number-13: 978-0-309-39252-5
International Standard Book Number-10: 0-309-39252-7
Digital Object Identifier: 10.17226/21921

Additional copies of this report are available for sale from the National Academies Press, 500 Fifth Street, NW, Keck 360, Washington, DC 20001; (800) 624-6242 or (202) 334-3313; http://www.nap.edu.

Printed in the United States of America

Suggested citation: National Academies of Sciences, Engineering, and Medicine. (2016). *Commercial Motor Vehicle Driver Fatigue, Long-Term Health, and Highway Safety: Research Needs.* Washington, DC: The National Academies Press. doi: 10.17226/21921.

The National Academies of
SCIENCES • ENGINEERING • MEDICINE

The **National Academy of Sciences** was established in 1863 by an Act of Congress, signed by President Lincoln, as a private, nongovernmental institution to advise the nation on issues related to science and technology. Members are elected by their peers for outstanding contributions to research. Dr. Ralph J. Cicerone is president.

The **National Academy of Engineering** was established in 1964 under the charter of the National Academy of Sciences to bring the practices of engineering to advising the nation. Members are elected by their peers for extraordinary contributions to engineering. Dr. C. D. Mote, Jr., is president.

The **National Academy of Medicine** (formerly the Institute of Medicine) was established in 1970 under the charter of the National Academy of Sciences to advise the nation on medical and health issues. Members are elected by their peers for distinguished contributions to medicine and health. Dr. Victor J. Dzau is president.

The three Academies work together as the **National Academies of Sciences, Engineering, and Medicine** to provide independent, objective analysis and advice to the nation and conduct other activities to solve complex problems and inform public policy decisions. The Academies also encourage education and research, recognize outstanding contributions to knowledge, and increase public understanding in matters of science, engineering, and medicine.

Learn more about the **National Academies of Sciences, Engineering, and Medicine** at **www.national-academies.org**.

PANEL ON RESEARCH METHODOLOGIES AND STATISTICAL APPROACHES TO UNDERSTANDING DRIVER FATIGUE FACTORS IN MOTOR CARRIER SAFETY AND DRIVER HEALTH

MATTHEW RIZZO (*Cochair*), Department of Neurological Sciences, University of Nebraska Medical Center
HAL S. STERN (*Cochair*), Department of Statistics, Donald Bren School of Information and Computer Sciences, University of California, Irvine
DANIEL BLOWER, University of Michigan Transportation Research Institute
CHARLES A. CZEISLER, Division of Sleep Medicine, Harvard Medical School, and Brigham and Women's Hospital
DAVID F. DINGES, Department of Psychiatry, University of Pennsylvania School of Medicine
JOEL B. GREENHOUSE, Department of Statistics, Carnegie Mellon University
FENG GUO, Department of Statistics, Virginia Tech Transportation Institute
RICHARD J. HANOWSKI, Center for Truck and Bus Safety, Virginia Tech Transportation Institute
NATALIE P. HARTENBAUM, Occumedix, Inc.
GERALD P. KRUEGER, Krueger Ergonomics Consultants
MELISSA M. MALLIS, M3 Alertness Management, LLC
JOHN R. PEARSON, Council of Deputy Ministers Responsible for Transportation and Highway Safety
DYLAN S. SMALL, Department of Statistics, The Wharton School, University of Pennsylvania
ELIZABETH A. STUART, Department of Mental Health and Biostatistics, Johns Hopkins Bloomberg School of Public Health
DAVID H. WEGMAN, Department of Work Environment, School of Health and Environment, University of Massachusetts, Lowell

MICHAEL L. COHEN, *Costudy Director*
ESHA SINHA, *Costudy Director*
RICHARD PAIN, *Consultant*
AGNES GASKIN, *Administrative Assistant*

COMMITTEE ON NATIONAL STATISTICS

LAWRENCE D. BROWN (*Chair*), Department of Statistics, The Wharton School, University of Pennsylvania
JOHN M. ABOWD, School of Industrial and Labor Relations, Cornell University
FRANCINE BLAU, Department of Economics, Cornell University
MARY ELLEN BOCK, Department of Statistics (emerita), Purdue University
MICHAEL CHERNEW, Department of Health Care Policy, Harvard Medical School
DONALD DILLMAN, Social and Economic Sciences Research Center, Washington State University
CONSTANTINE GATSONIS, Department of Biostatistics and Center for Statistical Sciences, Brown University
JAMES S. HOUSE, Survey Research Center, Institute for Social Research, University of Michigan
MICHAEL HOUT, Department of Sociology, New York University
THOMAS MESENBOURG, U.S. Census Bureau (retired)
SUSAN MURPHY, Department of Statistics and Institute for Social Research, University of Michigan
SARAH NUSSER, Office of the Vice President for Research, Iowa State University
COLM O'MUIRCHEARTAIGH, Harris School of Public Policy Studies, University of Chicago
RUTH PETERSON, Criminal Justice Research Center, Ohio State University
ROBERTO RIGOBON, Sloan School of Management, Massachusetts Institute of Technology
EDWARD SHORTLIFFE, Department of Biomedical Informatics, Columbia University and Arizona State University

CONSTANCE F. CITRO, *Director*
BRIAN HARRIS-KOJETIN, *Deputy Director*

BOARD ON HUMAN-SYSTEMS INTEGRATION

NANCY COOKE (*Chair*), Cognitive Engineering Research Institute, Arizona State University
ELLEN J. BASS, Department of Systems and Information Engineering, Drexel University
SARA J. CZAJA, Departments of Psychiatry and Behavioral Sciences and Industrial Engineering, University of Miami
FRANCIS T. DURSO, Department of Psychology, Georgia Institute of Technology
ANDREW S. IMADA, A.S. Imada & Associates, Carmichael, California
EDMOND ISRAELSKI, Human Factors Program, AbbVie, Abbott Park, Illinois
ELIZABETH LOFTUS, Criminology, Law and Society; Cognitive Sciences; School of Law, University of California, Irvine
FREDERICK L. OSWALD, Department of Psychology, Rice University
KARL S. PISTER, Department of Civil and Environmental Engineering, University of California, Santa Cruz
DAVID REMPEL, Division of Occupational Medicine, University of California, San Francisco
EMILIE M. ROTH, Roth Cognitive Engineering, Menlo Park, California
BARBARA SILVERSTEIN, Safety and Health Assessment and Research for Prevention Program, Washington State Department of Labor and Industries
DAVID H. WEGMAN, School of Health and Environment, University of Massachusetts, Lowell

POORNIMA MADHAVAN, *Director*

TRANSPORTATION RESEARCH BOARD
2016 EXECUTIVE COMMITTEE

JAMES M. CRITES (*Chair*), Dallas–Fort Worth International Airport, Texas
PAUL TROMBINO III (*Vice Chair*), Iowa Department of Transportation, Ames
NEIL J. PEDERSEN (*Executive Director*), Transportation Research Board

MEMBERS

VICTORIA A. ARROYO, Georgetown Climate Center, Centers and Institutes, and Environmental Law Program, Georgetown University Law Center, Washington, DC
SCOTT E. BENNETT, Arkansas State Highway and Transportation Department, Little Rock
JENNIFER COHAN, Delaware Department of Transportation, Dover
MALCOLM DOUGHERTY, California Department of Transportation, Sacramento
A. STEWART FOTHERINGHAM, School of Geographical Sciences and Urban Planning, Arizona State University, Tempe
JOHN S. HALIKOWSKI, Arizona Department of Transportation, Phoenix
MICHAEL W. HANCOCK, Kentucky Transportation Cabinet, Frankfort
SUSAN HANSON, Graduate School of Geography (emerita), Clark University, Worcester, Massachusetts
STEVE HEMINGER, Metropolitan Transportation Commission, Oakland, California
CHRIS T. HENDRICKSON, Department of Engineering, Carnegie Mellon University, Pittsburgh, Pennsylvania
JEFFREY D. HOLT, Power, Energy, and Infrastructure Group, BMO Capital Markets Corporation, New York
ROGER B. HUFF, HGLC, LLC, Farmington Hills, Michigan
GERALDINE KNATZ, Sol Price School of Public Policy, Viterbi School of Engineering, University of Southern California, Los Angeles
YSELA LLORT, Miami, Florida
JAMES P. REDEKER, Connecticut Department of Transportation, Newington
MARK L. ROSENBERG, The Task Force for Global Health, Inc., Decatur, Georgia
KUMARES C. SINHA, Department of Civil Engineering, Purdue University, West Lafayette, Indiana
DANIEL SPERLING, Department of Civil Engineering, Department of Environmental Science and Policy, and Institute of Transportation Studies, University of California, Davis

Copyright © National Academy of Sciences. All rights reserved.

KIRK T. STEUDLE, Michigan Department of Transportation, Lansing
GARY C. THOMAS, Dallas Area Rapid Transit, Texas
PAT THOMAS, State Government Affairs, UPS, Washington, DC
KATHERINE F. TURNBULL, Texas A&M Transportation Institute, College Station
DEAN WISE, Burlington Northern Santa Fe Railway, Fort Worth, Texas

EX OFFICIO

THOMAS P. BOSTICK, U.S. Army Corps of Engineers, Washington, DC
JAMES C. CARD, TRB Marine Board, The Woodlands, Texas
ALISON JANE CONWAY, Department of Civil Engineering, City College of New York, and TRB Young Members Council
T.F. SCOTT DARLING III, Federal Motor Carrier Safety Administration, U.S. Department of Transportation
MARIE THERESE DOMINGUEZ, Pipeline and Hazardous Materials Safety Administration, U.S. Department of Transportation
SARAH FEINBERG, Federal Railroad Administration, U.S. Department of Transportation
LEROY GISHI, Division of Transportation, Bureau of Indian Affairs, U.S. Department of the Interior, Washington, DC
JOHN T. GRAY II, Policy and Economics, Association of American Railroads, Washington, DC
MICHAEL P. HUERTA, Federal Aviation Administration, U.S. Department of Transportation
PAUL N. JAENICHEN, SR., Maritime Administration, U.S. Department of Transportation
THERESE W. MCMILLAN, Federal Transit Administration, U.S. Department of Transportation
MICHAEL P. MELANIPHY, American Public Transportation Association, Washington, DC
GREGORY G. NADEAU, Federal Highway Administration, U.S. Department of Transportation
MARK R. ROSEKIND, National Highway Traffic Safety Administration, U.S. Department of Transportation
CRAIG A. RUTLAND, U.S. Air Force Civil Engineer Center, Tyndall Air Force Base, Florida
REUBEN SARKAR, U.S. Department of Energy
BARRY R. WALLERSTEIN, South Coast Air Quality Management District, Diamond Bar, California
GREGORY D. WINFREE, Office of the Secretary, U.S. Department of Transportation

Copyright © National Academy of Sciences. All rights reserved.

FREDERICK G. (BUD) WRIGHT, American Association of State Highway and Transportation Officials, Washington, DC
PAUL F. ZUKUNFT, U.S. Coast Guard, U.S. Department of Homeland Security

Acknowledgments

The panel is grateful to the Federal Motor Carrier Safety Administration for providing the funds that made this study possible. Steven Smith, Office of Analysis, Research, and Technology, and his colleagues Albert Alvarez, Terri Hallquist, and Martin Walker, made presentations to the panel, provided us with technical reports, and answered many questions that arose during the course of the study. We especially want to thank Albert Alvarez, serving as the contracting officer's technical representative, who was extremely supportive of the panel's work.

The panel also appreciates those who provided important information for this study in presentations during our four information-gathering meetings: Paul Albert, Greg Belenky, Michael Belzer, Linda Boyle, Rebecca Brewster, Lamont Byrd, Jeff Dawson, Tom DiSalvi, Carol Flannagan, Rick George, Martial Hebert, Jeff Hickman, Paul Jovanis, Stephen Keppler, Elizabeth Klerman, Adrian Lund, Michael McDonal, Tony McDonald, David Marker, Daniel Mollicone, Christopher Monk, Dan Murray, Ryan Olson, David Osiecki, Don Osterberg, Robert Pentracosta, Stephen Popkin, Karl Sieber, Juna Snow, Tianjia Tang, Andrew Tarko, Matt Thiese, Pierre Thiffault, Tom Weakley, Rusty Weiss, and Ann Williamson. All of these individuals spent a great deal of their time preparing these presentations to advance the panel's work.

The panel also is indebted to staff of the National Academies of Sciences, Engineering, and Medicine outside of the Committee on National Statistics. They included staff in the Transportation Research Board, especially Steve Godwin, who provided highly useful advice on potential

panel members, Kenneth Campbell, and William Rogers. We also thank staff of the Division of Behavioral and Social Sciences and Education's Board on Behavioral, Cognitive, and Sensory Sciences, including its director, Barb Wanchisen, and Toby Warden.

The panel is very grateful as well to Ron Knipling for his comprehensive review of the research literature on fatigue and highway safety. We are also greatly in the debt of Rick Pain, who delayed his retirement to provide us with critical information and advice throughout all phases of this study. His contribution can be seen in every chapter.

Agnes Gaskin dealt with the complicated administrative aspects of such a study with great patience and understanding. Genie Grohman and Rona Briere did an outstanding job of technical editing of the report, greatly improving its readability, and helping to communicate our message.

Finally, we wish to think the panel members. The panel worked extremely well together with a great collaborative spirit, writing the majority of the chapters and conscientiously reviewing the products of other panel members. They were a wonderful group of people to get to know.

This report has been reviewed in draft form by individuals chosen for their diverse perspectives and technical expertise, in accordance with procedures approved by the Report Review Committee of the Academies. The purpose of this independent review is to provide candid and critical comments that will assist the institution in making its published report as sound as possible and to ensure that the report meets institutional standards for objectivity, evidence, and responsiveness to the charge. The review comments and draft manuscript remain confidential to protect the integrity of the process.

We wish to thank the following individuals for their review of this report: Paul S. Albert, Biostatistics and Bioinformatics Branch, Division of Intramural Population Health Research, Eunice Kennedy Shriver National Institute of Child Health and Human Development; Mathias Basner, Division of Sleep and Chronobiology, University of Pennsylvania Perelman School of Medicine; Michael H. Belzer, Department of Economics, Wayne State University; Stephen V. Burks, University of Minnesota, Morris; Paul P. Jovannis, Department of Civil and Environmental Engineering (emeritus), Pennsylvania State University; Stefanos N. Kales, T.H. Chan School of Public Health and Occupational Medicine Residency, Harvard School of Public Health; Tanzy M.T. Love, Department of Biostatistics and Computational Biology, University of Rochester Medical Center; Fred Mannering, University of South Florida; David Melton, retired, Global Road Safety, Liberty International; Daniel Mollicone, Pulsar Informatics, Inc., Seattle, Washington; Louis J. Ptacek, Division of Neurogenetics, Department of Neurology, University of California, San Francisco; Nina L. Shattuck, Human-Systems Integration Program, Operations Research

Department, Naval Postgraduate School; and Ann Williamson, Transport and Road Safety Research, School of Aviation, University of New South Wales. Although the reviewers listed above provided many constructive comments and suggestions, they were not asked to endorse the report's conclusions or recommendations, nor did they see the final draft of the report before its release. The review of this report was overseen by Richard W. Pew, Raytheon BBN Technologies, and James O. Berger, Department of Statistical Science, Duke University. Appointed by the Academies they were responsible for making certain that an independent examination of this report was carried out in accordance with institutional procedures and that all review comments were carefully considered. Responsibility for the final content of this report rests entirely with the authoring committee and the institution.

Matthew Rizzo and Hal S. Stern, *Cochairs*
Michael L. Cohen and Esha Sinha, *Costudy Directors*
Panel on Research Methodologies and Statistical
Approaches to Understanding Driver Fatigue Factors
in Motor Carrier Safety and Driver Health

Contents

PART I: BACKGROUND

PART II: CURRENT RESEARCH DATA AND METHODS

PART III: CURRENT RESEARCH FINDINGS

PART IV: RESEARCH DIRECTIONS

Summary

Approximately 4,000 fatalities result from crashes involving trucks and buses in the United States each year. Although estimates are wide-ranging, 10 to 20 percent of these crashes may have involved fatigued drivers. The stresses associated with work as a commercial motor vehicle (CMV) driver (e.g., irregular schedules, economic pressures) and the lifestyle many of these drivers lead put them at substantial risk for insufficient sleep and the development of short- and long-term health problems. Sleep disorders such as obstructive sleep apnea (OSA), for example, appear to be common among many CMV drivers. OSA is a major contributor to driver fatigue, which in turn raises a driver's risk for involvement in crashes. Moreover, it has become increasingly clear that drivers who regularly obtain insufficient sleep, whether as a result of irregular work patterns, sleep disorders, or other reasons, are likely at increased risk for a number of serious long-term health problems.

To address this problem, the Federal Motor Carrier Safety Administration (FMCSA) asked the National Academies of Sciences, Engineering, and Medicine to convene the Panel on Research Methodologies and Statistical Approaches to Understanding Driver Fatigue Factors in Motor Carrier Safety and Driver Health. The panel was charged with providing advice on additional data collection and analytic techniques with the potential to support a more comprehensive understanding of the links between operator fatigue and highway safety and between fatigue and long-term health problems, such as cardiovascular diseases. Specifically, the charge to the panel was to "assess the state of knowledge about the relationship of factors such as hours of driving, hours on duty, and periods of rest to the fatigue experienced by truck and bus drivers while driving and the implications for the safe operation of their vehicles. The panel will also assess the relationship of these factors to drivers' health over the longer term. It will identify improvements in data and research methods that can lead to better understanding in both areas."

STUDY CONTEXT

FMCSA's mission is to "reduce crashes, injuries, and fatalities involving large trucks and buses." The agency works toward this goal in at least three ways. First, a major policy lever at FMCSA's disposal is that it issues and administers hours-of-service (HOS) regulations for truck and bus drivers that specify the maximum number of hours they can work in a day and in a workweek. The hope is that if they drive limited hours, drivers will have enough time to obtain adequate sleep between work shifts, and therefore will be more alert and less fatigued while driving. As a result, it is believed, the risk of crashes will be lower.

Second, FMCSA is responsible for the medical certification of CMV drivers through the National Registry of Certified Medical Examiners (NRCME). Members of the NRCME examine CMV operators at least every 2 years to determine whether they meet FMCSA's medical standards.

Third, Transport Canada, FMCSA, trucking industry trade associations, and other agencies developed the North American Fatigue Management Program (NAFMP), an Internet-based online educational program that informs drivers, their employers, and anyone involved in commercial carrier operations about the causes of driver fatigue, the increased risk of crashes due to fatigue, the long-term health consequences of CMV driving (such as regular sleep insufficiency), and, most important, suggests countermeasures that can be used to manage driver fatigue.

The specifics of these three FMCSA programs are based on the current scientific understanding of operator fatigue, its causes, and its consequences. A considerable amount of research has been conducted to clarify the relationship among hours of service, driver fatigue, and crash risk, as well as the relationship between fatigue and long-term health and wellness. As yet, knowledge of these relationships is not comprehensive, and the relationships themselves may be changing. However, the quality and quantity of the information available on these relationships is constantly increasing. In particular, new in-vehicle technologies and roadway improvements are regularly being applied to promote safety, and several of these innovations allow for the capture of important data that could be used to better inform policy and/or improve driving procedures. Additionally, recent advances in statistical methods could be applied to make better practical and research use of data that either are currently collected or could be collected if targeted for research purposes.

HOS regulations can only limit hours spent driving and working; they cannot mandate rest, so they inherently cannot ensure, by themselves, that drivers will be well rested and alert. Therefore, it is not straightforward to determine how additional modifications of the current HOS regulations would result in more or less fatigue in CMV drivers that might, respectively, raise or lower crash risk.

In addition, driver fatigue obviously is not the only cause of highway crashes. Crash risk factors can be grouped into at least four main types: driver characteristics; truck or bus characteristics; factors stemming from employment circumstances, especially with respect to scheduling and work assignments; and the physical environments encountered while driving. The variety of factors from which a crash can result, acting either solely or in combination, makes it challenging to develop an understanding of the nature of an individual risk factor—in the present case fatigue—since there are so many confounders that are difficult to control or otherwise account for in analyses. Therefore, the study designs and analytic tools used in such research are critically important.

A further complication is that fatigue is very difficult to define and therefore to measure objectively. If fatigue is loosely defined as the inability to sustain performance over time, under such a vague definition, it is not directly measurable. Therefore, it is somewhat difficult to assess fatigue, and thus to regulate how to avoid driving while fatigued. As a result, researchers and policy makers must instead assess how hours of service affect various components of fatigue, many of which are measurable. One of the most important, objectively measureable components of fatigue is "drowsiness" or lack of alertness.

To add to this complexity, the commercial truck and bus industries are highly heterogeneous, with a great variety of types of employment, methods for compensation, and so on. As a result of their specific type of employment, many CMV drivers are affected in different ways, and sometimes not at all, by changes in HOS regulations. Therefore, stratifying, or otherwise accounting for, the type of employment in some way is important when carrying out research in this area.

Adding still further to this complexity is the difficulty of assessing some of the primary inputs—sleep duration, hours of service, level of driver alertness—and the primary output—rates of crashes due to fatigue. Deriving satisfactory measures of the amount or quality of sleep obtained by a truck or bus driver during the previous night or nights is difficult because self-reports often are not highly reliable and because attempts to directly capture measures of drivers' sleep are invasive in nature. Also, it has been shown that paper logs recording hours of service for CMV drivers are often inaccurate. In crash reporting databases, moreover, the primary judges of the factors involved in a highway crash are police officers, who must make such assessments after the fact, with little information to go on. Partly as a result of these difficulties, the available crash reporting databases often provide a paucity of information on sleep deficiency in CMV drivers, their adherence to HOS regulations, and their crash frequency as a result of fatigued driving. Therefore, research on the linkage among hours of service, fatigue, and accident frequency is hampered by imperfect knowledge of the three most central variables.

THE RESEARCH QUESTIONS MOST ESSENTIAL TO FMCSA

Any newly proposed changes to the HOS regulations for CMV drivers, to the process for drivers' medical certification, and to the NAFMP need to be based on research-supported understanding of the costs and benefits of such changes. The following list represents the panel's attempt to articulate questions that, if answered satisfactorily, should assist FMCSA in understanding the costs and benefits of existing and proposed changes to its policies and regulations in these areas:

- How much sleep do typical CMV drivers need to maintain suitable sustained levels of alertness and to avoid being drowsy to the point of driving while impaired?
- To what extent would any proposed change in HOS regulations affect the amount of sleep obtained by CMV drivers in different industry sectors?
- What degree of hypopnea (severity level of OSA) results in enough sleep loss to increase the risk of crashes for CMV drivers?
- To what extent does regular use of positive airway pressure (PAP) and related OSA treatment technologies and measures mitigate that increased risk?
- To what extent are various collision avoidance and driver fatigue alert technologies (both in-vehicle technologies and infrastructure measures such as roadway rumble strips) useful for reducing the risk of crashes?
- What substances, if any, reduce impairment due to sleep insufficiency?
- To what extent is chronic sleep deprivation related to an increased risk of developing health threats or various medical conditions?

- To what extent do CMV drivers, their employers, corporate officials, fleet supervisors, safety and risk managers, and drivers' families make use of the NAFMP materials on the Internet?
- To what extent do fatigue awareness training and fatigue management initiatives result in behavioral improvements in CMV drivers?

Considerable progress has been made toward answering many of these questions using techniques ranging from laboratory observation and driving simulators for focused comparisons; to various kinds of epidemiological techniques, especially case-control and cohort studies; to naturalistic driving studies assessing what happens in the field. Laboratory studies can enable comparison of specific treatments or interventions in highly specialized environments; however, the necessary extrapolation from laboratory observation to field implementation is not always straightforward. Epidemiological studies and naturalistic driving studies, and even summary crash data, can be highly informative about what is happening in the field; however, these studies can be subject to confounding influences when attempts are made to compare interventions or treatments.

SUMMARY OF WHAT IS KNOWN

What Causes Driver Fatigue?

Fatigue-related performance decrements are a complex function of several factors, including lengthy time on task (i.e., long drives), extended wakefulness or acute sleep deprivation, chronic insufficient sleep, and poor-quality sleep. Some of these factors are attributable to irregular work schedules, nighttime work, and misalignment of circadian phase. Other possible contributors to driver fatigue include the presence of sleep disorders, such as OSA, and other medical disorders, and even use of sedating medications. However, the extent to which each of the above individual factors, or the interaction among them, adversely affects performance is unknown.

Why Do Drivers Continue to Drive When They Are at Risk of Fatigue?

Drivers get inadequate sleep for several reasons. CMV driving is an essential part of the just-in-time delivery economy, so the demand for such workers is increasing. To understand why CMV drivers may continue to drive when sufficiently drowsy to be impaired, one must consider the form of their work compensation structure and many other confounding influences, including family pressures and commuting patterns. The factors involved in decision making are relatively well known but not their specific individual contributions.

What Is the Relationship Between Sleep Deficits and Decreased Driver Alertness?

Research has shown that insufficient sleep leads to decreased alertness and eventually to performance decrements. Performance decrements can lead in turn to driver errors or inappropriate driving practices, which then can lead to crashes. These errors and inappropriate practices usually are a function of slowed reaction times, attention failures, and poor decision making.

What Are Practicable Ways of Reducing or Eliminating Driver Fatigue?

Various fatigue countermeasures are commonly attempted, some that have been tried and proven to work, and others that yield only fleeting effects and are not as helpful. Among those that work best are adhering to work-rest scheduling that permits sufficient sleep, driving primarily during the daytime rather than at night, being cognizant of the two anticipated circadian lulls of the 24-hour day, obtaining sleep immediately prior to a long trip, planning to take and taking periodic breaks from driving during trips, and inserting planned naps into a trip plan. While consuming caffeine can provide temporary relief, and rumble strips can serve to alert drivers that they are likely falling asleep, these measures really provide only temporary assistance. The only way to reduce the need for sleep before going on duty and thus to alleviate or prevent operator fatigue, aside from short and temporary postponements, is to obtain an adequate quality and quantity of sleep.

What Is the Relationship Between Acute or Chronic Sleep Loss and Increased Crash Risk for CMV Drivers?

Acute fatigue is thought to develop when drivers shortchange themselves of several hours of sleep per 24-hour day over a few days in succession or even over a workweek. Such acute fatigue can usually be ameliorated by gaining sufficient recovery sleep (e.g., during a driver's nominal 2 days off or his/her weekend). On the other hand, *chronic fatigue* may be more long-lasting and, in addition to repeated sleep shortages, may involve elements of continuous pressure or stresses from other sources (e.g., job or family and home circumstances). Knowledge currently is lacking about how to identify chronically fatigued drivers and how to determine ways in which drivers and employers can deal with the problem.

What Is the Relationship Among OSA, PAP Use, and Crash Risk?

OSA is associated with increased crash risk for CMV drivers. An important need is to determine how to identify those drivers most at risk of crashes due to OSA. The extent to which treatment of severe OSA with PAP devices mitigates the risks due to OSA is unknown, but there is some evidence that PAP use for 4 or more hours a night brings the risk for crashes close to that for drivers without OSA. Also unknown is whether drivers diagnosed with mild or moderate OSA are at significant crash risk or whether they, too, might benefit from PAP treatment to make them more alert while driving.

How Do Health and Wellness Programs and Fatigue Management Initiatives Modify the Behavior of CMV Drivers?

A few large truck carriers have derived positive results from their almost 10 years of experience in integrating health and wellness and fatigue management programs, and they have shared those experiences, including the return on their investment in such initiatives. However, most studies of these programs have not sufficiently and reliably validated their efficacy for achieving the goal of reducing crash risk or their scalability. Also, little is known about the use of health and wellness programs by independent owner-operators. Additional development, research, and demonstration of validity and reliability are needed in this area.

What Are the Likely Safety Effects of Improvements in and Deployment of Collision Avoidance and Fatigue Alert Technologies?

Studies of the potential of improved collision avoidance and fatigue alert technologies to ameliorate the impact of fatigue on crash risk in a variety of situations have been conducted in laboratories, in driving simulators, in the field, and in demonstrations of newer technological innovations. In general, however, such studies have not sufficiently and reliably validated the efficacy of these technologies in reducing crash risk or their readiness for deployment. Additional development, research, and demonstration of validity and reliability are needed in this area.

DIRECTIONS FOR FUTURE RESEARCH

Inputs and Outcomes Essential for Further Research Progress

Further progress in research on the relationship among CMV driver fatigue, drivers' long-term health, and highway safety will require quality information on the inputs and outputs corresponding to the essential research questions listed above. Specifically, information is clearly needed on the following:

- driver loss of alertness in near real time (possibly obtained by a system that measures degree of eyelid closure);
- amount of quality sleep a driver received in the past 24 hours;
- the number of sleep-related crashes in which a driver was involved per vehicle-miles traveled over a long period of time (e.g., several years of driving);
- a driver's development of various health conditions that can affect alertness;
- changes in components of a driver's lifestyle, especially diet and health; and
- a driver's number of lane deviations, unusual speed changes, and unusual brake applications.

Since roadway crashes are rare events, one measurement outcome often used as a surrogate for crashes in naturalistic driving studies is safety-critical events (SCEs). SCEs often are defined as kinematic events, such as hard braking or swerving, that can be viewed as actions that nearly resulted in crashes and are therefore reasonable surrogates for those events. SCEs include incidents that are and are not fatigue related. For some research purposes, then, only a subset of SCEs is relevant, so methods for identifying the most relevant subset of SCEs for research on CMV driver fatigue need to be determined.

To learn more about the research questions outlined above, it will be necessary to better understand the links between the relevant inputs and the relevant outcomes. Achieving this understanding will in turn require having these input and outcome variables at the level of the individual driver. And obtaining that information will likely require the ability to merge data across separate sources. At present, however, while the inputs or the outputs for many of the above research questions may be available to some parties (e.g., confidentially held by carriers/employers or accessibly electronically recorded by equipment manufacturers), many of these measures are not available to researchers.

Alternative Sources of Data

A large proportion, but not all, of the data needed to answer important research questions either is currently collected or could be collected in the near future. These data might be collected by various government agencies, by individual trucking or busing companies, by companies that help monitor the functioning of drivers and vehicles, by industrial collectives, and by various electronic recording devices on the vehicles being driven. In particular, much of these data are currently held by individual companies or industrial collectives with a promise of confidentiality that precludes their sharing for any purpose, including research. This information might include, for example, data on crashes per vehicle-mile traveled, video data of the driver that could be used to assess PERCLOS (percentage of eye closure),[1] evidence of lane weaving and hard braking, information on the driving environment and the condition of the vehicle, maneuvers made immediately prior to crashes and in advance of noncrashes, demographics, driver height and weight, other driver health measures, and the type of employment. The collection of these data by private companies and their collectives demonstrates the feasibility of gathering such information and may suggest alternative data collection approaches for FMCSA to consider.

Furthermore, several data-handling procedures and statistical techniques have been used over the past 25 years to provide protection against the release of individually identifiable information but may not be well known outside of the various national statistical agencies. Use of such techniques might make it possible to revise some confidentiality agreements to allow for sharing of selected information with researchers.

For many of the research questions posed above, one potential drawback is that even after the various inputs and outcomes had been gathered, one would have to link those variables for individual drivers and vehicles (properly anonymized). Development of such linkages might require the assignment of a code that would identify a driver across data sets.

Finally, since some of the linkages and information on long-term driver health outcomes will be very difficult to acquire, research progress in this area would be advanced if the National Institute for Occupational Safety and Health (NIOSH), in conjunction with the Bureau of Labor Statistics and FMCSA, were to design and carry out an ongoing survey on the health and wellness of CMV drivers. A survey that would allow comparisons over time is needed to better understand the dynamics underlying changes in the health and wellness of truck and bus drivers.

---

[1] A measure of the percentage of eyelid closure over a period of time (one of the most accepted measures of drowsiness).

Use of Causal Inference Techniques

Crashes often have many causes, only one of which may be driver fatigue due to sleep deficiency, and these other causes must be accounted for in research aimed at understanding how fatigue is linked to crashes. Further, since identifying the causes of a serious crash after the fact is difficult, one must be aware of the potential for substantial measurement error. Research therefore is needed that can disentangle various individual inputs, such as fatigue, from other confounding factors with respect to their impact on crash risk. Such research will require the use of recently developed data collection and statistical techniques useful for drawing causal inferences. It would be advisable for FMCSA to promote the use of such techniques in the research it supports.

Validation of New Technologies

A number of technological countermeasures have been and are being introduced (e.g., lane-deviation tracking systems, newer-generation PERCLOS tracking systems) with the promise of helping to mitigate driver fatigue. The National Highway Traffic Safety Administration could fill an important need by outlining a procedure that could be followed to validate the effectiveness of such new technologies.

Validation of Educational Programs

To date, the effectiveness of the NAFMP has not been properly assessed. The panel believes strongly that the longitudinal survey on CMV drivers' health and welfare recommended below (in Recommendation 10) needs to include questions on interaction with the NAFMP.

KEY CONCLUSIONS

CMV Drivers' Health and Wellness

Conclusion 1: Insufficient sleep can increase the risk of developing various health problems, including obesity, diabetes, hypertension, and cardiovascular disease, all of which can impact an operator's level of alertness while driving and potentially impact crash risk.

Conclusion 2: Based on the evidence on drivers who are not commercial motor vehicle drivers, obstructive sleep apnea is known to increase crash risk, and there is no evidence base or compelling reason for thinking *that the same would not also be true among commercial motor vehicle drivers.*

Conclusion 3: Better understanding is needed of the effects of treating obstructive sleep apnea in commercial motor vehicle drivers with positive airway pressure (PAP) therapy with respect to the amount and quality of sleep they obtain and their cognition and driver performance following PAP treatment sessions.

Conclusion 5: Substantial data gaps limit understanding of factors that impact the health and wellness of commercial motor vehicle drivers. Closing these gaps would aid greatly in developing a better understanding of drivers' current status and long-term prospects with respect to health and wellness.

Conclusion 8: Insufficient information exists on (1) how the variety of fatigue management and health and wellness management programs available have been designed, (2) whether drivers/employers actually adhere to these programs, and (3) whether these programs are effective in achieving their goals.

Sleep Insufficiency and Its Impact on Highway Safety

Conclusion 10: There is no biological substitute for sufficient sleep.

Conclusion 12: Despite almost three decades of research on the topic, technological innovations for detecting driver fatigue in near real time and operational strategies for their use are still in the early phases of understanding and application.

Conclusion 13: Biomathematical models can be useful for the development of general work-rest schedules. However, existing models do not account for individual variation, so care must be taken in applying them to address likely impacts of irregular work schedules.

RECOMMENDATIONS

Data Gaps

RECOMMENDATION 1: The National Institute for Occupational Safety and Health should be enlisted to design and conduct a regularly scheduled survey every 5 to 10 years to gather information

needed to better understand the demographics and employment circumstances of all commercial motor vehicle drivers in various industry segments.

RECOMMENDATION 2: The Federal Motor Carrier Safety Administration should conduct an evaluation to determine whether commercial motor vehicle drivers' use of electronic on-board recorders correlates with reduced frequency of hours-of-service violations and reduced frequency of crashes compared with those drivers who do not use such instruments.

RECOMMENDATION 3: Given the potential research benefits of the use of data from electronic logging devices, Congress should consider modifying Title 49 of the U.S. Code to permit the use of such data for research purposes in a manner that protects individualized confidential data from disclosure, and if such a change is made, the Federal Motor Carrier Safety Administration should make parallel provisions in its regulations.[2]

RECOMMENDATION 4: When commercial trucks and buses containing electronic data recorders that record data on the functioning of the driver and the truck or bus are involved in serious crashes, the relevant data should be made available to investigators and to safety researchers.

RECOMMENDATION 5: The Federal Motor Carrier Safety Administration should incentivize those that capture driver performance data (e.g., large fleets, independent trucking associations, companies that collect telematics data, insurance companies, researchers) to increase the availability of those data relevant to research issues of operator fatigue, hours of service, and highway safety. Any such efforts should ensure that data confidentiality is maintained, perhaps through restricted access arrangements or use of statistical techniques for disclosure protection.

RECOMMENDATION 6: The Federal Motor Carrier Safety Administration should work to improve the collection of and/or access to baseline data on driving exposure by including in its data collection efforts greater detail on the driving environment and by providing these data at low levels of geographic aggregation—even for indi-

---

[2] A change has been made from the prepublication copy to update language to make it clear that FMCSA cannot change the law but it can modify its regulations.

vidual highway segments. Comparisons enabled by the availability of these baseline data would benefit several proposed lines of new research.

How CMV Drivers React to Feelings of Drowsiness

RECOMMENDATION 7: The Federal Motor Carrier Safety Administration should support research aimed at better understanding the factors associated with driver behavior related to fatigue and sleep deficiency, including what motivates drivers' decisions about whether to continue driving when they feel fatigued.

Testing of New Technologies

RECOMMENDATION 8: Using a human-systems integration framework, the Federal Motor Carrier Safety Administration and the National Highway Traffic Safety Administration, in consultation with the Centers for Disease Control and Prevention and the National Institutes of Health, should develop evaluation guidelines and protocols for third-party testing, including field testing, conducted to evaluate new technologies that purport to reduce the impact of fatigue on driver safety.

How FMCSA Can Improve Its Research Projects Through Peer Review

RECOMMENDATION 9: The Federal Motor Carrier Safety Administration should make greater use of independent peer review in crafting requests for proposals, assisting in decisions regarding awards, and monitoring the progress of projects (including in the study design and analysis stages). Peer review should include expertise from all relevant fields, including epidemiology and statistics—especially causal inference—to address appropriate design and analysis methods.

Understanding of What Impacts the Long-Term Health of CMV Drivers

RECOMMENDATION 10: The U.S. Department of Health and Human Services and/or the U.S. Department of Transportation should fund, design, and conduct an ongoing survey that will allow longitudinal comparisons of commercial motor vehicle drivers to enable tracking of changes in their health status, and the factors

likely to be associated with those changes, over time. In addition, it would be highly desirable for the survey data thus collected to include sufficient information to enable linking of the data to relevant electronic health records, with a particular focus on conditions that may threaten drivers' health and safety.

RECOMMENDATION 11: The Federal Motor Carrier Safety Administration should continue to encourage all individuals included in the National Registry of Certified Medical Examiners to utilize current best practices in identifying drivers who should be referred for additional sleep malady testing and in making determinations about commercial driver's license renewal extensions. It would be highly preferable, as soon as possible, to supply the examiners with clear criteria or guidance on when it is appropriate to refer presenting drivers for sleep malady testing.

RECOMMENDATION 12: The Federal Motor Carrier Safety Administration should support peer-reviewed research on obstructive sleep apnea (OSA) and commercial motor vehicle drivers throughout all the research stages, from the drafting of requests for proposals through analysis of data. The supported research should be focused on a better understanding of the incidence of OSA in commercial motor vehicle drivers; its impact on driver fatigue, safety, and health; and the benefits of treatments. Specific research topics might include

- **determining the number of commercial motor vehicle drivers whose quantity/quality of sleep and driving performance are likely affected at various levels of apnea-hypopnea (index of OSA severity);**
- **determining what rules for sleep-screening referrals are effective in discriminating between those commercial motor vehicle drivers with and without OSA;**
- **delineating the causal chain from diagnosis of OSA (preferably as a function of severity) to increased likelihood of crash frequency among commercial motor vehicle drivers;**
- **determining the impact of treatment with positive airway pressure (PAP) and similar devices on long-term health and crash rates among commercial motor vehicle drivers with varying degrees of apnea severity; and**
- **identifying the required/recommended duration of initial PAP treatment (e.g., suggested number of hours of treatment per day/week) before a driver can be certified to return to driving.**

RECOMMENDATION 13: The Federal Motor Carrier Safety Administration (FCMSA) should carry out a research program on driver fatigue management and training. This research program should include

- **evaluating the effectiveness of the North American Fatigue Management Program (NAFMP) for educating truck and bus drivers in how to modify their behavior to remedy various potential sources of fatigue;**
- **determining how effective the NAFMP training modules are in meeting the needs of drivers' employers, including fleet managers, safety and risk managers, dispatchers, driver trainers and other corporate officials (e.g., those conducting carrier-sponsored employee health and wellness programs);**
- **evaluating any new education programs regarding sleep apnea that FMCSA has or plans to develop; and**
- **examining possibilities for the development and evaluation of incentive-based programs for improving health and fitness, including regular coaching, assessment, and support.**

1

Introduction

Approximately 4,000 fatalities due to truck and bus crashes occur each year, up to 20 percent of which are estimated to involve fatigued drivers. Commercial motor vehicle (CMV) drivers may decide to drive when they are fatigued as a result of either external demands (e.g., delivery schedules) or internal decisions (e.g., a need to get to the desired location). Consequently, the Federal Motor Carrier Safety Administration (FMCSA), an agency of the U.S. Department of Transportation, has established hours-of-service (HOS) regulations that limit the number of hours CMV drivers can drive and the number of hours in which they can engage in other tasks, such as loading and unloading, on both a daily and weekly basis. While HOS regulations are based on a substantial research literature on how they impact fatigue, fatigue and HOS regulations are two among many possible causal factors implicated in crash risk. A better understanding and mitigation of the risk posed by fatigue in commercial driving could be obtained through the acquisition of more relevant data and the use of more targeted quantitative methods.

To help answer questions about the linkages among hours of service, fatigue, highway safety, and the long-term health of CMV drivers, FMCSA requested that the National Academies of Sciences, Engineering, and Medicine, through its Committee on National Statistics, convene the Panel on Research Methodologies and Statistical Approaches to Understanding Driver Fatigue Factors in Motor Carrier Safety and Driver Health; see Box 1-1 for the panel's detailed statement of task. This request came at an opportune time because many new and anticipated data sources on motor

BOX 1-1
Statement of Task

An ad hoc panel will conduct a study to assess the state of knowledge about the relationship of factors such as hours of driving, hours on duty, and periods of rest to the fatigue experienced by truck and bus drivers while driving and the implications for the safe operation of their vehicles. The panel will also assess the relationship of these factors to drivers' health over the longer term. It will identify improvements in data and research methods that can lead to better understanding in both areas. The study is requested by the Federal Motor Carrier Safety Administration (FMCSA) in the U.S. Department of Transportation, which is responsible for regulating hours of service (HOS), including time on duty and time and periods of driving, for motor carrier operators engaged in interstate commerce. The panel's review will cover a broad range of literature on factors that relate to fatigue, impaired performance, and adverse health outcomes for motor carrier operators and workers in other industries that involve similar cognitive and physiological demands, including but not limited to analyses and modeling studies carried out by FMCSA. Based on its review and deliberations, the panel will issue a report with findings and recommendations that (1) assesses the strength of the evidence (based on the quality of the research methods and the underlying data) regarding factors, such as hours of driving, hours of duty, and periods of rest, that may lead to fatigue and impaired cognitive and physiological performance of motor carrier drivers; (2) assesses the strength of the evidence regarding these factors and impaired health outcomes (including effects on mortality and morbidity) for motor carrier operators; (3) identifies priorities for research and modeling to improve knowledge of HOS and other factors in motor carrier driver fatigue, safety, and health; and (4) identifies the most promising data collection methods (e.g., anonymous response surveys, naturalistic driving studies, electronic on-board recorders, other fatigue management technologies) and the most appropriate statistical methods for analyzing very large data sets (e.g., data mining, modeling, generation of synthetic data) to support state-of-the-art research in these important areas.

The panel will take account of the regulatory context that underlies FMCSA's interest in research and data collection on factors, including hours of service and periods of rest, that can affect motor carriers' cognitive and physiological performance on the job and their longer term health. However, it will not recommend changes to HOS rules.

vehicle crashes and their causes either are now available or soon will be. Examples include on-board electronic recording systems and video data capture and vehicle telematics, often associated with naturalistic driving studies. In addition, to help analyze the more detailed and complex data derived from these new sources, new statistical and computer science methodologies have been developed. In particular, methods for assessing causal relationships have been improved substantially in the past 20 years. Techniques for analyzing data with a complex time and spatial correlation structure, including mixed linear models, also have been developed and are relevant to the analysis of data in this area.

STUDY CONTEXT

Fatigue and Performance in Safety-Sensitivity Occupations

HOS regulations in regulated industries typically specify both the limits of time at work and minimum rest time in order to maintain safety and reduce risk resulting from fatigue. *Fatigue* generally refers to a subjective sense of weariness, but in work environments, it refers primarily to the objective decline in performance resulting from physical exertion and/or behavioral effort over time, as well as from inadequate time for recovery. The concept of degraded performance with time on task has its origins in the physical meaning of fatigue relative to biology (e.g., continued stimulation leading to temporary loss of responsivity) and physics (e.g., increasing tendency of a material to break down after repeated physical stress). When applied to human performance, fatigue refers to increasing performance variability and instability in behavioral alertness and vigilance due to continued time on task without breaks (Lim et al., 2010) and/or during night work (Neri et al., 2002) and/or after sleep loss (Basner and Dinges, 2011; Lim and Dinges, 2008). These effects are captured in the definition of fatigue as "a biological drive for recuperative rest" (Williamson et al., 2011).

The contribution of inadequate sleep to fatigued performance in safety-sensitive occupations has been of increasing concern as a result of mounting scientific evidence over the past 20 years that obtaining sufficient sleep is essential for optimal brain function and performance (Banks and Dinges, 2007; Lim and Dinges, 2008; Xie et al., 2013; Yang et al., 2014). Population studies reveal that many people experience insufficient sleep for medical and occupational reasons (Basner et al., 2014; Ford et al., 2015). The U.S. Centers for Disease Control and Prevention (CDC) reports that the number of American adults sleeping 6 hours or less per day increased from 38.6 million to 70.1 million between 1985 and 2012, an increase the CDC characterizes as a "public health epidemic" (Centers for Disease Control and Prevention, 2015; Ford et al., 2015). An estimated 37 percent of the adult U.S. population sleeps less than 7 hours per day, which recent evidenced-based consensus reviews and epidemiological studies have found to be the minimum sleep duration to prevent cumulative deterioration in performance and health and increased risk of mortality (Ferrie et al., 2007; Watson et al., 2015a, 2015b). Studies suggest that the majority of short sleepers do not require less sleep than other adults;

rather, these individuals gradually accrue sleep debt over time and report a greater tendency to fall asleep unintentionally (McKnight-Eily et al., 2011; Punjabi et al., 2003). Individuals who obtain less sleep than needed because of work typically sleep longer on nonwork days—a pattern seen in CMV drivers (Dinges et al., 2005b; Hanowski et al., 2007) that suggests many accrue a sleep debt on workdays.

Fatigue is a demonstrated risk to safety in transport and occupational settings (Williamson et al., 2011). In addition to compromising alertness and vigilance, prolonged time on task, inadequate breaks, reduced sleep time, and limited time off duty, individually and collectively, can slow reaction time and cognitive processing speed and degrade working memory, situational awareness, and impulse control. The result is an increase in both performance errors of omission, which involve the failure to respond in a timely manner, and errors of commission, which involve responding prematurely or incorrectly (Lim and Dinges, 2010). Subjective ratings of fatigue generally underestimate the magnitude of performance deficits due to fatigue (Banks et al., 2010; Van Dongen et al., 2003a), which limits the utility of self-reported fatigue relative to safety risk.

In current commercial transportation systems, the most common causes of fatigue relate to the interaction of eight temporal domains: (1) amount of time awake before work; (2) time of day work occurs (i.e., night versus day); (3) amount of time at work (i.e., on-duty duration); (4) amount of time working specific tasks (e.g., driving versus not driving); (5) time for acute rest during work periods (e.g., number and duration of breaks); (6) time not working (i.e., off-duty duration within and between days); (7) daily time asleep (i.e., both acute and cumulative); and (8) degree of sleep disruption (e.g., by untreated sleep apnea, pain, and other factors). HOS typically involve limitations in a subset of these temporal domains.

Because the concept of fatigue was first associated with deteriorating performance due to time on task, before the critical contributions of sleep need and circadian timing were understood, time working (on duty) and time on task (driving) have received much of the attention relative to fatigue mitigation through HOS regulations for CMV drivers. However, scientific studies in the past 25 years have established that driver fatigue and performance also are dynamically influenced by the regulation of sleep need and endogenous circadian rhythms, including the need to obtain sufficient sleep to ensure recovery from work schedules that might induce either acute or chronic sleep deprivation. The major advance in understanding of work-related fatigue has derived from scientific evidence that the neurobiology underlying human performance is such that changes in performance are not simple functions of amount of sleep, circadian cycle, and other factors.

Fatigue and Commercial Motor Vehicle Operators

An example of this nonlinearity is illustrated by the findings of a naturalistic study of CMV drivers. This study found that approximately 30 percent of all observed instances of driver drowsiness occurred within the first hour of the work shift, and that drowsiness was twice as likely to occur between 6 AM and 9 AM compared with baseline or nondrowsy driving (Barr et al., 2011). This finding appears to be the opposite of the expected adverse effects of driving time on alertness. However, a database study of all fall-asleep crashes in North Carolina found that they peaked in frequency between 6 AM and 9 AM (Pack et al., 1995). Also important, scientific studies of the interaction of sleep and circadian dynamics have established that these are among the hours of the day when sleep propensity is especially elevated as a result of little or no sleep the night before (Lim and Dinges, 2008), after repeated days of sleep restriction (Cohen et al., 2010; Mollicone et al., 2010), or during sleep inertia from just having awakened (Jewett et al., 1999). Therefore drowsy driving would be expected to be higher at 6-9 AM than at 6-9 PM (the circadian peak in daily alertness), even though 6-9 AM represents the beginning hours of a drive (Burke et al., 2015). Thus, an understanding of fatigue-related accidents and errors among CMV drivers needs to be based on the best available information on how time on task, time awake, sleep time, and circadian time interact relative to risk, in addition to well-recognized contributors to risk (e.g., road conditions, traffic density, environmental factors).

Fatigue interacts with other aspects of the lifestyles of CMV drivers, including unhealthy diet and insufficient exercise. A substantial body of evidence indicates that a chronic reduction in sleep time—especially to 6 or fewer hours per day, which has been objectively documented among CMV drivers (Dinges et al., 2005b; Hanowski et al., 2007; Van Dongen and Mollicone, 2014)—is associated with many long-term health problems, including obesity, hypertension, diabetes, and cardiovascular disease, as well as performance deficits (Watson et al., 2015a, 2015b). Evidence that obesity is the avenue by which the other disorders occur was found in a recent study of 88,246 CMV medical examinations from 2005 to 2012. The cohort had a 53-percent prevalence of obesity (defined as body mass index [BMI] > 30.0 kg/m^2), and obesity was linked to heart disease, hypertension, diabetes mellitus, nervous disorders, sleep disorders, and chronic low back pain in these individuals (Thiese et al., 2015a). This study also found that between 2005 and 2012, the prevalence of morbid obesity (defined as BMI > 35.0 kg/m^2) increased 8.9 percent, and the prevalence of three or more co-occurring medical conditions associated with obesity that limit driving certification increased fourfold.

Obesity and related health conditions are of concern not only for driver health in general but also for driver retention and for the increased risk of fatigued driving. As a result, FMCSA joined with other U.S. and Canadian agencies to develop the North American Fatigue Management Program (NAFMP), an online information and educational program (described in Chapter 8) designed to help drivers understand the factors that contribute to their fatigue and its consequences and to encourage healthier habits.[1] In 2010, the National Transportation Safety Board (NTSB)[2] made the following recommendation:

> To be most effective, a fatigue management program should be comprehensive and authoritative. Within the next 2 years, the NAFMP is expected to provide fatigue management program guidelines specifically designed for use in the motor carrier environment. Implementation of these guidelines by every motor carrier would be a major step toward addressing the problem of fatigue among commercial drivers on the nation's highways. But if the NAFMP guidelines remain voluntary—and are used by some carriers but ignored by others—this important safety tool might have only a limited effect in reducing fatigue-related highway accidents. Consequently, the NTSB recommends that the FMCSA require all motor carriers to adopt a fatigue management program based on the NAFMP guidelines for the management of fatigue in a motor carrier operating environment.

However, the effectiveness of the NAFMP among CMV drivers is unknown, as is the case for alternative approaches to fatigue management. Consequently, the NTSB stated ". . . FMCSA, as the Federal agency responsible for motor carrier safety, must also be involved in the evaluation of the fatigue management programs used by carriers to determine whether they successfully mitigate fatigue." The NTSB concluded that if fatigue management programs are to be successful, FMCSA oversight is needed; therefore, the NTSB made the following recommendation to FMCSA: "Develop and use a methodology that will continually assess the effectiveness of the fatigue management plans implemented by motor carriers, including their ability to improve sleep and alertness, mitigate performance errors, and prevent incidents and accidents."

DATA LIMITATIONS

Upon reviewing the data available for developing an understanding of how driver fatigue relates to crash risk, the panel identified a number

---

[1] Available: http://www.nafmp.com/en [March 2016].

[2] National Transportation Safety Board Safety Recommendation Date: October 21, 2010. In reply refer to H-10-8 through -11 and H-08-13 and -14 (Reiteration).

of limitations. First, any analysis of fatigue, hours of service, and crash frequency must reflect the substantial heterogeneity in the types of jobs performed by CMV drivers. Differences exist in the size of the fleets with which they are affiliated, truck-handling characteristics, whether trucks are monitored electronically, drivers' work schedules, how drivers are compensated, the length of time drivers are away from home, and other factors. Although considerable data are collected on drivers who work for large carriers, much less information is available on those who work for small carriers and, especially, on independent owner-operators.

Second, the data available to researchers on various causal factors are a patchwork, with some of the most essential variables being either not recorded, imperfectly captured, or recorded but without the information needed to evaluate their linkages to crash risk.

Third, the statistical methods used in analyzing the available data often fail to take adequate account of confounding influences. More extensive and sophisticated analysis methods are needed because analysis of the data sets now available, as well as those being collected in even larger naturalistic studies, needs to account for the effects of a wide range of factors, in addition to and in interaction with fatigue, to determine the extent to which they affect crash risk among CMV drivers. These factors can be grouped broadly into the characteristics of the driver (which include the temporal domains relevant to fatigue, as discussed above), the carrier, the driving environment, and the vehicle. Given the complexity of many types of confounding factors, different outcomes of interest (e.g., fatal crashes or all crashes), different subsets of the populations of CMV drivers, and different types of jobs, it is not surprising that efforts to date to examine the causal association among driver behavior, hours driving, driver fatigue, and crash rates have resulted in a variety of somewhat disparate findings.

Better quality data, including temporal synchronization across variables, and the use of advanced quantitative methods could provide an understanding of the multiple interacting factors that increase crash risk. This understanding would in turn inform future HOS regulations and the development of other types of countermeasures, and help ensure that they instantiate a more comprehensive picture of the relationship between driver fatigue and highway safety.

FMCSA also is responsible for overseeing the regular medical certification of CMV drivers through the efforts of medical examiners who certify that they are fit to drive. These examinations take place at least every 2 years and are intended to evaluate the risk of sudden or gradual impairment or incapacitation due to medical conditions or their treatment. Evidence on the effectiveness of these medical exams in evaluating drivers who are likely to have sleep apnea and those who are susceptible to fatigue for other medical reasons is limited. Some evidence does indicate that, since the National Registry of Certified Medical Examiners was implemented, examiners have been inconsistent in the criteria they apply in evaluating the risk of medical conditions that may lead to driver fatigue. This inconsistency represents another avenue of opportunity for identifying ways to improve CMV drivers' health and highway safety.

ORGANIZATION OF THE REPORT

Part I of this report provides background information on the problem of CMV driver fatigue and its relationship to drivers' long-term health and highway safety. Chapter 2 describes the trucking and bus industries. Chapter 3 presents what is known about fatigue, operator performance, and safety. Chapter 4 then describes the HOS regulations and provides a short history and some international comparisons.

Part II of the report describes the relevant available data and the relevant newer methodologies for analyzing those data. Chapter 5 describes the data available from a variety of sources, while Chapter 6 details the types of methods that could be applied to analysis of the data.

Part III of the report provides an in-depth discussion of what is currently known about fatigue. Chapter 7 addresses the effects of fatigue on highway safety. Chapter 8 describes how fatigue relates to health. Chapter 9 describes countermeasures for dealing with the effects of fatigue. Chapters 8 and 9 also present the panel's conclusions.

Part IV addresses research directions. Chapter 10 describes the research needed on fatigue and highway safety. Chapter 11 describes the research needed on the role of fatigue in long-term health and on management of fatigue. These two chapters present the panel's recommendations.

Finally, it should be noted that, because of an imbalance in the available literature and knowledge, as well as the difference in the degree of prevalence of the problem, this report is somewhat more targeted to truck than to bus safety. However, both issues are covered in the chapters that follow.

PART I

BACKGROUND

2

The Trucking and Bus Industries

Truck and bus operations in the United States are as diverse as the U.S. economy. Considerable heterogeneity in the truck and bus industries stems from operational characteristics ranging from fleet size and employer type to work schedules and on-the-job activities. Given this diversity, it is difficult to make simple statements about the relationship of such factors as work hours and periods of rest to fatigue among commercial motor vehicle (CMV) drivers and about drivers' long-term health. Nonetheless, to help the reader understand the population regulated by the Federal Motor Carrier Safety Administration (FMCSA), this chapter reviews characteristics of the U.S. trucking and bus industries and provides a brief description of the lifestyles of CMV drivers and the policies and practices that influence driver fatigue and health. The chapter also touches on the attitudes in both industries toward driver fatigue and health and wellness programs and the information available on demographic and anthropometric variables with regard to CMV drivers.

Transportation in the United States occurs by road, rail, air, and waterways (boats) and can be broadly categorized into passenger and freight transportation. Passenger traffic is handled primarily by private automobiles (80%), planes (12%), trains (1%), and buses (7%) (U.S. Department of Transportation, 2014, Table 2-1). Freight traffic provides U.S. households with access to manufactured goods, and business establishments rely on freight transportation to move their raw materials and finished products. Trucks carry 67 percent of U.S. freight by weight and a much higher percentage by value. Trucks are crucial not only in the transportation industry but also in many other sectors, such as agriculture, construction, and warehousing.

THE TRUCKING INDUSTRY

The trucking industry comprises hundreds of thousands of carriers and millions of drivers moving goods locally or in long hauls between cities. The industry is diverse, and different segments have different operational characteristics. FMCSA regulates trucks that are operated in interstate commerce or the transport of hazardous materials in quantities requiring a placard. There currently exists no single method for classifying truck drivers, but the major domains of classification are highlighted below.

Size of Carrier

Important differences exist among drivers by virtue of the size of the carrier for which they drive. Drivers may be employed by a large fleet or a small fleet, or they may be a one-truck operation, that is, an independent owner-operator.[1] Table 2-1 shows the distribution of carriers and trucks by fleet size, displaying clearly the majority-minority industry structure. The majority of trucks on the road are operated by large carriers even though they represent a minority of the number of carriers in the industry. A very few of the large carriers are quite large, owning thousands of trucks. Conversely, as the table shows, smaller firms—those with 20 or fewer trucks—make up an appreciable fraction of the industry: nearly 95 percent of the more than 500,000 active truck carriers.

Types of Trucks

As shown in Table 2-2, truck drivers operate a variety of types of trucks. As a result, driver licensing requirements differ depending on the type of truck, as well as the type of load (discussed below). Any driver of a truck with a 26,001 or more lb gross vehicle weight rating (GVWR) must have a valid commercial driver's license (CDL). The requirements for a valid CDL are set nationally, while each state regulates CDL testing and issues the licenses. Endorsements to the CDL are required for certain categories of truck and load type. For example, drivers of multiple-trailer trucks or any truck carrying hazardous materials must have the appropri-

---

[1] There is no standard definition for "large" or "small" fleet. In compliance reviews published by FMCSA, carrier fleet size is categorized as very small (1-6 power units), small (7-20 power units), medium (21-100 power units), and large (> 100 power units). Available: https://ai.fmcsa.dot.gov/SafetyProgram/spRptReview.aspx?rpt=RVFS [March 2016].

TABLE 2-1 Distribution of Carriers and Trucks as of November 2014

Fleet Size (based on number of trucks)	Carriers	% of Total	Number of Trucks	% of Total
Between 1 and 3	397,328	72.91	584,923	13.48
Between 4 and 20	119,148	21.86	922,497	21.26
Between 21 and 55	16,247	2.98	529,354	12.20
Between 56 and 100	4,185	0.77	310,197	7.15
Between 101 and 999	3,849	0.71	967,986	22.31
Greater than/equal to 1,000	283	0.05	1,023,681	23.59
Zero	409	0.08	0	0.00
Unknown	3,478	0.64	0	0.00
TOTAL	**544,927**	**100.00**	**4,338,638**	**100.00**

NOTE: The companies are all registered with the Federal Motor Carrier Safety Administration to operate trucks in interstate commerce or to transport hazmat, and they have "recent activity."
SOURCE: Truck count based on Motor Carrier Management Information System Carrier File.

TABLE 2-2 Classes of Trucks Operated by Truck Drivers by Gross Vehicle Weight Rating

GVWR Class	GVWR Range (lbs)	Example	Type	Driver Needs CDL?
3	10,001 to 14,000	Big pickup, delivery van	Straight	No
4	14,001 to 16,000	Delivery van	Straight	No
5	16,001 to 19,500	Small dump truck, medium van	Straight	No
6	19,501 to 26,000	Utility truck, home fuel oil delivery truck	Straight	No
7	26,001 to 33,000	Dump truck, big-box van, 2-axle tractor (used with two 28.5-foot trailers)	Mainly straights, two-axle tractors	Yes
8	33,001 and above	3-axle tractor, big dump truck, concrete mixer, 3-axle straight van	Mainly tractors, 3+-axle straights	Yes

NOTES: GVWR = gross vehicle weight rating, the maximum operating weight of a vehicle, including the driver and the cargo, as specified by the manufacturer. CDL = commercial driver's license, which permits one to operate a commercial motor vehicle of a specific size.
SOURCE: For greater detail, see Burks et al. (2010).

ate endorsement. To qualify for this endorsement, they must demonstrate additional knowledge and skills beyond those required for the basic CDL. Since each state issues CDLs, FMCSA, in concert with the states, maintains a national database of CDL licenses, so a driver should not be able to obtain multiple licenses, and a driver with a suspended or revoked license in one state is precluded from continuing to drive a truck with a license from a different state.

Following are the common truck types driven by truck drivers:

- two- and three-axle straight trucks, with van, dump, or tank cargo bodies, such as the two-axle, six-tire vans used for local package delivery;
- two- and three-axle straight trucks pulling a trailer, which include small vans or dump trucks with equipment trailers;
- bobtail tractors (tractors with no trailer);
- tractor-semitrailers; and
- tractor two- or three-trailer combinations, which include the so-called STAA[2] doubles (two 28-foot trailers), turnpike doubles (two 40+-foot trailers), and triple-trailer combinations.

Employer

Many truck drivers are employed by private carriers that are not for-hire trucking firms. Private carriers transport their own goods or use the trucks in furtherance of their business. This means the operations of the trucks (times, road types, loads, etc.) are constrained by the needs of the business. Some private carriers move such a large volume of freight that they can operate like a big truckload (TL) or less-than-truckload (LTL) for-hire carrier (see definitions of TL and LTL in the following subsection). These are manufacturers, distributors, or retailers that move their own goods among factories, distribution centers, and retail outlets. They can have regular schedules, routes, and driving times, which can be controlled to address driver needs for regular rest, returning home every night after work, and so on. The drivers' jobs can be like any other regular shift jobs. Other private carriers use trucks incidental to their primary business, such as construction, landscape, or retail.

Unlike the private carriers described above, for-hire carriers primarily provide freight transportation services for their customers; their business is not to transport their own freight.

---

[2] STAA refers to the Surface Transportation Assistance Act of 1982, which allows large trucks, referred to as STAA trucks, to operate on routes that are part of the National Network.

Type of Load

Within the for-hire carrier domain, the distinction between TL and LTL services can have important ramifications for driver schedules and rest. In the case of TL services, the load is typically being transported from a single shipper to a single destination. LTL services combine shipments from multiple shippers and transport them to multiple destinations.

LTL carriers usually aggregate multiple small shipments into a load, transport them between terminals (as described below), and then distribute the shipments locally. TL carriers generally do not have such a terminal structure and focus on transporting large shipments, often for relatively long distances. Drivers for TL carriers typically pick up a full load from a shipper and move it directly to the receiver of the goods. A large majority of TL carriers' business is regular and predictable under contracts or less formal agreements. Some TL carriers provide dedicated service, regularly hauling loads for a particular customer. Drivers in these situations may have predictable and regular schedules.

Big TL firms move enough freight that they can control dispatch to allow relatively regular schedules for their drivers. Other types of TL services are more occasional and episodic. For example, a TL firm or independent owner-operator may work with a freight broker. Such brokers find loads and arrange for carriers or owner-operators to haul them. In these situations, drivers can have much less regular schedules and unpredictable routes. The driver picks up a load, delivers it, then checks in with the broker for the next load; overall this process can result in unpredictable and irregular schedules.[3]

In contrast, a defining feature of LTL service is that, as noted above, many small loads from a variety of shippers are consolidated into one load (at a terminal), transported to destination terminals, and then dispersed to their destinations. Some LTL drivers operate primarily terminal to terminal, on regular schedules and routes. They drive tractors pulling one, two, or three trailers. At the terminals, the consolidated load may be broken down into smaller loads and then loaded on medium-duty vans and delivered by other drivers to recipients. Thus some LTL drivers make long-haul trips in big tractor-trailer rigs, while others perform package delivery in smaller trucks, in neighborhoods and to businesses. Although the package delivery drivers operate within a local area, their routes may vary day by day. They also spend a good part of each day loading and unloading at multiple locations.

Finally, some truck drivers work for private carriers, for which, as

---

[3] On the other hand, owner-operators in this situation are their own bosses and can more easily take a break when needed, without having to get permission or be accountable to a dispatcher.

described above, the truck operation is incidental to the main line of business. Some big retailers operate their own trucking arms, which provide LTL-like service between suppliers and retail stores. At the other end of the spectrum are construction firms that use large trucks to transport building materials and farmers who haul agricultural products to market.

Relationship to Hours-of-Service Regulations and Driver Fatigue

The major domains of classification described above influence drivers' work schedules, the types and lengths of routes they drive, and what constitutes their daily duty. Given the scope of this report, it is useful to describe the complexities of these influences from the perspective of hours-of-service (HOS) regulations and driver fatigue:

- Regular versus irregular schedules
    - LTL service typically entails more regular schedules, both for long-haul and package delivery operations. Regular schedules make it easier for drivers to comply with HOS regulations and may give them more opportunities to get regular rest and sleep at home each night.
    - Private carriers may have regular schedules for the most part, with variations for some industries and in the case of some seasons or circumstances. For example, farmers and custom harvesters operate in response to the needs dictated by crops, while utility workers may work extended hours to restore service following an outage.
    - Large TL carriers generally can provide more regular schedules and hauls, although that may not always be the case. Smaller TL carriers are likely to have less regular schedules and hauls, and their drivers may be away from home for longer periods and have less control over their schedules.
    - Hazmat drivers transport dangerous cargo for which a timely delivery schedule becomes primary, which may place the driver in conflict with HOS regulations.
- Regular versus irregular routes
    - As described above, TL carriers may be more likely than LTL carriers to have irregular routes, which in turn create difficulties in adhering to HOS regulations. Driver fatigue may be an issue when schedules are irregular.[4]

---

[4] Crum and colleagues (2002) investigated the influence of motor carrier scheduling practices on driver fatigue for 116 truck companies and 66 motor coach companies and found that regular schedules and routes appeared to diminish fatigue.

  - — LTL package delivery entails elements of irregular routes, but the delivery operations are typically on a regular schedule, which should be conformable to the limits set by HOS regulations.
  - — Private carriers generally have regular routes and can manage operations to maintain regular schedules. As noted above, there are exceptions depending on their operations (e.g., seasonal demands related to agriculture or utility work).
- Length of haul[5]
  - — LTL and TL carriers, irrespective of fleet size, are more likely to have long than shorter routes.
  - — Large private carriers, such as the big retail companies, are more likely to have long than shorter routes.
  - — Smaller private carriers are more likely to have local or regional routes than longer routes.
- Primary activity in the job
  - — In for-hire operations, driving is the job almost by definition, so for-hire truck drivers are more likely than drivers for private carriers to face difficulties related to the limits of the HOS regulations, and driving-related fatigue may be an issue for them.
  - — There is more variability in private operations.
    - o In some private operations, the primary job of the driver is driving. For these drivers, HOS regulations and driving-related fatigue are salient but fairly readily controlled in a well-managed firm.
    - o In other private operations, truck driving is incidental to the job. HOS, especially time behind the wheel and driving-related fatigue, are less of an issue for these drivers unless they are told to put in more hours driving to get a critical job done. Under such circumstances, HOS regulations become relevant.
- Loading/unloading
  - — Big LTL carriers are involved in terminal-to-terminal operations; drivers likely are not involved in loading/unloading of their cargo.
  - — Drivers working for TL carriers, particularly small ones, may load and unload their cargo.
  - — The same may be true for drivers for private firms, except those working for big retail distribution companies.

---

[5] The terms *long haul*, *regional*, and *short haul* are often used to indicate the typical distance traveled to deliver a load. Long haul often means greater than 700 miles, regional between 300 and 700 miles, and short haul less than 300 miles.

— Drivers for small private farms and similar businesses are likely to load and unload trucks in addition to driving. Driving is not their sole job by definition.

Given the heterogeneity of the operational structure of truck driving, it is no surprise that big carriers have a different cost structure from that of medium-sized carriers and independent owner-operators—each industry segment faces a different average cost and average benefit curve. The market is highly segmented, with multiple players and intense competition leading to thin profit margins. According to *IBIS World*, the top six carriers represented about 10 percent of total sales in 2012. There are a small number of large players and a large number of small players, and none of the players hold substantial market power. As different players operate differently, the impact of HOS regulations varies across the industry. The challenge FMCSA faces is devising HOS regulations that do not compromise either the economic feasibility of different-sized trucking operations or highway safety.

WORKING CONDITIONS AND PAY OF A TRUCK DRIVER

Individuals working as long-haul truck drivers for smaller firms are more likely than drivers for larger firms to be away from home for most of the year and to be alone while driving. Their job is physically demanding because they may be required to load and unload cargo, which increases the risk of injuries. Given the current HOS regulations, being on duty 14 hours per day is common (Hege et al., 2015). Drivers can face long waiting periods during pickup and drop-off operations at warehouses and ports. The combination of long working hours, physically demanding tasks, and the pressures of a just-in-time economy creates a work environment for truck drivers that is not favorable to their health and has safety implications. Box 2-1 presents information from three studies on the workday of truck drivers.

Along with such issues, compensation provided to truck drivers plays a role in their decision to drive long hours. Truck drivers are paid in one of three general ways: (1) by the hour, (2) by the mile, or (3) by the load (which can include a percentage of the revenue associated with the load). The per-mile rate varies among employers and may depend on the type of cargo and the driver's experience. Some long-distance drivers, especially owner-operators, are paid a share of the revenue from shipping. Along with these three types of compensation schemes, there exist hybrids that entail some combination of base pay and incentives. According to the Bureau of Labor Statistics, the median annual wage for heavy and tractor-trailer truck drivers was $38,200 in May 2012, with a wide range: the low-

BOX 2-1
The Workday of Long-Haul Truck Drivers

Information from three studies provides some sense of the work schedule of a typical truck driver.[a]

Hanowski and colleagues (2000) collected information on activities performed by commercial motor vehicle (CMV) drivers in long/short-haul operations using a survey questionnaire. The 42 drivers sampled in the study spent about 28 percent of their workday driving, 35 percent loading/unloading, 26 percent on other assignments (e.g., merchandising, checking in/out, vehicle inspection), 7 percent waiting to unload, 2 percent eating, 0.5 percent resting, and 1.5 percent on other activities.

Subsequent work done by Soccolich and colleagues (2013) on the CMV driver workday using naturalistic data found that the majority of the 96 long-haul/line-haul drivers in the study sample spent their 14-hour workday driving (65.7%), followed by light work such as paperwork (18.7%); resting, which included sleeping and eating (11.5%); and doing heavy work such as loading/unloading (4.1%).

Jovanis and colleagues (2011) identified 10 common multiday (1-week) driving patterns of drivers based on paper logs[b] supplied by two big truckload (TL) and three less-than-truckload (LTL) carriers (Jovanis et al., 2011). The driving patterns showed significant variation in time of day when drivers began or ended their shifts, days of the week they were off duty, and hours devoted to sleeping/resting. A certain percentage of drivers were driving mainly from midnight to noon in the first 4 days of a week, taking off the next 2 days, and then returning to work on the last day of the week. Another group of drivers were on duty infrequently the first 4 days of the week and worked intensively the remaining days, with driving being concentrated from noon to midnight.

---

[a] The panel did not have comprehensive information on the work schedules of truck drivers; therefore, the information in this box should not be considered representative of the truck driver population.

[b] To ensure compliance with hours-of-service regulations, CMV drivers are required to fill out logs detailing their working hours, including time spent driving, on duty/not driving, off duty, and in a sleeper berth.

est 10 percent earned less than $25,110, while the top 10 percent earned more than $58,910 (U.S. Bureau of Labor Statistics, 2014).

Two general dimensions of driver compensation are relevant to this study:[6]

1. Level of compensation—offering higher pay gives carriers the opportunity to attract and retain more highly skilled and safer drivers.

---

[6] Key among the laws that impact the compensation of truck drivers is the Fair Labor Standards Act exemption issued in 1938, which exempted the trucking industry from overtime compensation. For details, see http://www.dol.gov/whd/flsa/ [March 2016].

2. Type of compensation—paying by the mile, by the hour, or by the trip encourages drivers to drive more miles, hours, or trips, which contributes to driver fatigue. Regardless of the compensation method, it is common practice not to pay for loading and unloading. The laborious nature of these activities contributes to driver fatigue, which may be exacerbated by the fact that drivers may try to rush through them given that the time is unpaid.

It is not clear how the compensation scheme and hours of service interact and which predominates in terms of influencing safe driving. Paying by the hour incentivizes drivers to accumulate more time. On non-Interstate/freeway-quality roads, driving time often increases because of traffic volume, lower speed limits, and traffic lights. Thus distance may decrease, but driving time increases. This puts compensation and HOS regulations in conflict.

Belzer (2000) makes the point that deregulation in the trucking industry has led to negative externalities—declining wages and poor working conditions. Deregulation led to extreme competition in certain sectors of the industry, which in turn drove down wages as carriers were keen to find clients, and driver wages were a primary target for cutting costs (Belzer, 2000).

Evidence suggests that compensation and safety are linked. Using driver-level data from a large carrier—J.B. Hunt—Rodriguez and colleagues (2006) highlight the results of an experiment done by the carrier when it increased driver wages by 39.1 percent (on average) in a single move. Controlling for demographic factors, work experience, and operational factors, the higher wages reduced driver turnover and improved safety performance (lower crash probability on average). Another study (Thompson et al., 2015) found that manipulation of payment methods in a simulator produced poor decision making in fatigued drivers.

THE BUS INDUSTRY

Table 2-3 lists total bus registrations in 2013 by type of ownership. Like the trucking industry, the bus industry has various segments. Based on the type of operation run, the bus industry can be divided into four sectors:

1. Scheduled service—Bus companies offering a posted service, such as that operated by Greyhound and Megabus, which runs specific routes regionally throughout the United States and Canada. Drivers bid on these routes, and successful drivers receive a 90-day contract.
2. Tour and charter—Bus companies specializing in group tours and

TABLE 2-3 Bus Registrations in 2013, by Type of Ownership

Ownership Type	Number
Private and commercial	137,656
School and other publicly owned	217,735
Federal	6,550
State, county, and municipal	496,572

SOURCE: U.S. Federal Highway Administration (2013, Table MV-10).

excursions. For this service, there is usually no repetitive fixed or specific route. Drivers are assigned this work a few days in advance.

3. Hybrid service—Bus companies operating both scheduled and tour and charter services.
4. Transit service—Buses operated by or for municipalities or regional authorities. They have a structured schedule by route.

Bus types and bus operators are at least as diverse as is the case for trucks, possibly more so. The above four are the dominant for-hire passenger operations, but there are others, including the following:

- school buses;
- shuttle buses belonging to any industry sector required to transport people as part of its service (e.g., airports, hotels, retirement homes, canoe liveries);[7]
- buses operated by private companies that transport employees between work sites or work locations;
- buses operated by municipalities and units of government that transport clients, prisoners, and workers; and
- buses operated by churches and other social service organizations.

Types of Buses

Types of buses driven in the United States include the following:

- School buses are specifically designed to transport school students.
- Motorcoaches are operated by bus companies running scheduled

---

[7] If a shuttle bus is big enough (seating for nine or more occupants, including the driver), it falls under HOS regulations.

and tour and charter services and are used to convey passengers between cities.

- Transit buses are used to provide local transportation services.
- Van-based buses, such as shuttle buses, are used by companies or organizations for transporting people.

Employer

Most transit bus drivers work for local governments or urban transit systems, which are private companies that contract with a city or town to provide bus service. Most charter bus drivers work in the charter bus industry. Intercity bus drivers typically work in the interurban and rural bus transportation industry. School bus drivers and special-client bus drivers are usually employed by a school district or private transportation company that contracts with a district to provide bus service. Some school bus services are provided by the local government.

Size of Fleet

As in the trucking industry, the firm size of bus carriers varies, from large intercity carriers that operate a fleet of buses to small charter/tour carriers operated as a family business. Although not large in number, some large school bus operators provide service under contract with school districts, while some school districts operate their own buses.

WORKING CONDITIONS AND PAY OF A BUS DRIVER

According to the Bureau of Labor Statistics, about one-half of all bus drivers worked full-time in 2012. The rest either worked part-time or had variable schedules. Most school bus drivers work only when school is in session. Some make multiple runs if the schools in their district open and close at different times. Others make only two runs, one in the morning and one in the afternoon, so their work hours are limited. Transit drivers may work weekends, late nights, and early mornings. Charter or tour bus drivers travel with their vacationing passengers. Driver hours are dictated by a tour schedule, and drivers may work all hours of the day, as well as weekends and holidays. Some intercity bus drivers have long-distance routes, so they spend some nights away. Other intercity bus drivers make a round trip and return home at the end of each shift.

The median annual wage for transit and intercity bus drivers, which includes charter bus drivers, was $36,600 in May 2012. The lowest-paid 10 percent earned less than $21,320, and the highest-paid 10 percent earned more than $59,480. The corresponding median annual wage of school or special client bus drivers was $28,080, with the lowest-paid 10 percent earning less than $17,610 and the highest-paid 10 percent earning more than $43,560.

ATTITUDES IN THE TRUCKING AND BUS INDUSTRIES TOWARD FATIGUE AND HEALTH AND WELLNESS PROGRAMS

The trucking and bus industries are playing a significant role in moving freight and people across the country, but safety remains an issue. In 2013, there were 3,541 fatal crashes involving large trucks in which 3,964 people were killed. The analogous statistics for fatal crashes involving a bus in 2013 are 280 and 310, respectively (U.S. Department of Transportation, 2013). Beyond loss of life, fatal crashes result in economic losses to the truck and bus companies and to society more generally.

Driver fatigue is one of many factors that contribute to crashes. How serious is the problem of fatigue in the bus and trucking industry? Estimates from police accident report-based databases, such as the General Estimates System and the Fatality Analysis Reporting System (FARS), suggest that fatigue plays a role in 1 to 2 percent of fatal crashes involving trucks and buses annually. However, these estimates are generally known to be underestimates because driver fatigue is difficult to detect. In 1990, the National Transportation Safety Board conducted in-depth investigations of 182 fatal-to-driver large-truck crashes and found fatigue to be a principal cause in 31 percent of these cases (56 of 182 crashes) (National Transportation Safety Board, 1990). Among all the factors investigated, fatigue turned out to be the most common cause. This 31-percent statistic should not be generalized to larger crash populations, such as “all fatal truck crashes” or “all truck crashes.” Nonetheless, it shows the importance of fatigue-related crashes for commercial drivers.

Putcha and colleagues (2002) analyzed FARS data looking for fatal crashes in which bus drivers were involved during the period 1995-1999. The five cases they found that were attributed to bus driver fatigue occurred in 1997, 1998, and 1999. In percentage terms, 0.3 percent of the bus drivers involved in fatal crashes during the 5-year period were coded as drowsy, fatigued, or asleep, which can be considered an underestimate for the reason mentioned above. One of the first studies of fatigue among bus drivers was conducted by Mackie and Miller (1978). Their most significant finding was that bus drivers who operate on irregular schedules suffer greater subjective fatigue and physiological stress relative to drivers operating on a regular schedule. A further description of studies investigating the linkage between driver fatigue and highway safety, including their limitations, is included in Chapter 7.

Efforts have been made over time to educate stakeholders in the trucking and bus industries on the issue of driver fatigue and the importance of the health and wellness of CMV drivers. An outreach and communication program initiated in 1996, a collaborative effort of the Federal Highway Administration's Office of Motor Carriers and Highway Safety and the American Trucking Associations, educated drivers and truck and bus companies about the risks of and the countermeasures for driver fatigue. The program arrived at a crucial finding: "if a driver's lifestyle could be focused on health, wellness, and fitness, it would be a precursor to overall safety consciousness" (Krueger, 2010b, p. 3). Efforts by other agencies, such as the Centers for Disease Control and Prevention and the National Institute for Occupational Health and Safety (NIOSH), either through research studies or conferences, have likewise made the point that the occupational health of CMV drivers is a central issue in highway safety.

It can safely be said that there is industry-wide acceptance that fatigue is an issue. Big trucking fleets take proactive approaches to improve their safety figures by formulating various health and wellness programs based on internal research and evaluation. Regarding bus companies, Greyhound has a fatigue management initiative based on input received from focus groups made up of drivers, operations managers, and safety directors throughout the industry. Independent owner-operators are left to their own devices and the North American Fatigue Management Program serves as an information source for them. Chapter 8 describes the various initiatives of particular carriers, FMCSA, and industry associations with regard to fatigue management and health and wellness programs. Chapter 11 addresses how to evaluate the effectiveness of a web-based approach to education.

DEMOGRAPHIC AND HEALTH INFORMATION ON TRUCK AND BUS DRIVERS

As noted in Chapter 1, demographic and anthropometric information on truck and bus drivers is needed to examine the relationship between hours of service (or hours of sleep) and crash frequency. In examining this relationship, various confounders must be kept in mind. For example, the relationship could be considerably different for young and old drivers. Also, it is important to know various medical and psychological characteristics of the drivers, including whether they suffer from medical conditions (e.g., obstructive sleep apnea [OSA]); whether they take prescribed medicines, many of which can make one drowsy; and whether they have a propensity for risky driving.

Obtaining demographic information on the truck driver population is difficult as there have been no censuses of truck drivers, and the surveys that have been conducted have by their design either excluded or underrepresented portions of the universe of truck drivers. Considerable information is available on drivers for most of the large carrier fleets. Also, the Owner-Operator Independent Drivers Association (OOIDA) has been conducting a biennial survey of its members since 1998. As is clear from its name, OOIDA focuses on owner-operators, so the information it collects excludes drivers that are employed by carriers. The drivers on whom these data are collected operate mainly tractor-semitrailers, which are almost all three-axle or larger trucks. The most salient feature of OOIDA members is that they are small carriers and probably operate a large number of TL loads.

According to OOIDA's 2012 membership profile, 97 percent of owner-operators are male, their mean age is 55, and 48 percent live in rural areas with populations of less than 5,000. Their average height is 5′11″, their average weight is 221 pounds, and their average body mass index (BMI) is 31. Twenty-nine percent have no medical insurance. Their mean net income is $55,367. They drive an average of 110,000 miles per year. Forty-seven percent are away from home more than 200 nights per year. Twenty-nine percent believe their medical plan covers sleep apnea, and 6 percent are receiving treatment for that condition (OOIDA Foundation, 2015). As seen in Table 2-1, there were more than 500,000 active truck carriers as of October 2014. Carriers with 1 to 3 trucks make up about 75 percent of motor carriers but operate only 13 percent of all trucks, while those motor carriers with 50 or more trucks account for 50 percent of all trucks. Also, there are some very large carriers with thousands of trucks, and while they are small in number, they employ a large number of truckers. Therefore, the demographic information from the OOIDA membership profile provides information for a relatively small percentage of truck drivers.

With respect to anthropometric information on truck drivers, some information on the general health of truck drivers was collected in a survey sponsored in late 2010 by NIOSH and carried out by Westat. The sample design randomly selected 1,670 truck drivers at randomly selected truck stops along randomly selected limited-access highway segments. Truck drivers were defined as people for whom truck driving was their main occupation. In addition, this survey was restricted to those driving trucks with three or more axles. Furthermore, the drivers had to have driven a heavy truck 12 months or longer and had to have taken at least one mandatory 10-hour rest period away from home during each delivery run. Eligibility was established in a preliminary interview, and 1,265 drivers completed the full survey. The findings indicated that 69 percent were obese, with 17 percent of all respondents being morbidly obese. Fifty-one percent smoked, 14 percent were diabetic, and 38 percent were not covered by health insurance. Twenty-seven percent admitted to sleeping on average less than 6 hours a night.

With regard to information on sleep received by truck and bus drivers, survey estimates are available from a poll of transportation workers conducted by the National Sleep Foundation in 2012 to understand their sleep habits and work performance. The poll was conducted among pilots, train operators, truck drivers, bus/limousine/taxi drivers, and a control group. Among the respondents were 203 truck drivers (mainly short-haul drivers employed by a company) and 116 bus drivers. Of these, 19 percent of truck drivers and 18 percent of bus/limousine/taxi drivers reported getting less sleep than needed.[8] These figures were considerably below what was reported by pilots and train operators. As noted earlier, not getting adequate sleep at night causes daytime sleepiness, which results in taking naps during workdays. Fully 42 percent of truck drivers and 53 percent of bus/limousine/taxi drivers reported taking naps on workdays. The average number of naps per day (average amount of time napping) taken by truck drivers was 3.4 (43.5 minutes) and by bus/limousine/taxi drivers was 3.5 (42.1 minutes). The survey also included items on sleep disorders. Eleven percent of truck drivers and 10 percent of bus/limousine/taxi drivers reported ever being diagnosed with a sleep disorder, with the majority (more than 80%) reporting that disorder to be OSA.

---

[8] Respondents were asked to compare the hours of sleep they needed according to survey workers with the hours of sleep they were actually getting on workdays. On the basis of the responses, the proportions of those getting (1) more sleep than needed, (2) sufficient sleep, and (3) less sleep than needed on workdays were calculated.

3

Consequences of Fatigue from Insufficient Sleep

The effect of fatigue on commercial motor vehicle (CMV) drivers is an important public safety issue. The National Transportation Safety Board has identified fatigue as a probable cause or a contributing factor in incidents and accidents across all modes of transportation. It is important to understand how fatigue affects performance and the implications of these effects for highway safety.[1]

Safe motor vehicle operation requires, among other things, the ability to stay awake and sustain maintenance of stable vigilance, situational awareness, and appropriately timed psychomotor and cognitive responses. However, these are the neurobehavioral functions affected most immediately and profoundly by work-hour fatigue and insufficient sleep. As noted in Chapter 1, fatigue refers to increasing performance variability, instability in behavioral alertness, and decreasing vigilance due to continued time on task without breaks (Lim et al., 2010; Neri et al., 2002) or insufficient sleep (Basner and Dinges, 2011; Lim and Dinges, 2008). The contribution of insufficient sleep to performance deficits in CMV driving is a reasonable concern given the scientific evidence that obtaining sufficient sleep is essential for optimal performance of tasks central to operating a motor vehicle safely (Banks and Dinges, 2007; Lim and Dinges, 2008, 2010).

---

[1] This chapter deals with fatigue and not distraction, which is also a common reason for highway crashes. Fatigue and distraction can be interactive. For more information on distraction, see Thiffault (2011).

Population studies have found that Americans reduce their sleep time primarily in favor of work time, commute time, and leisure time, in that order of importance (Basner and Dinges, 2009; Basner et al., 2007, 2014). As noted in Chapter 1, insufficient sleep is currently defined as sleeping less than 7 hours per day (Ford et al., 2015; Watson et al., 2015a, 2015b). This number derives from two sources. The first is randomized controlled laboratory experiments that systematically varied the duration of sleep obtained by healthy adults every day for 1 to 2 weeks, from 3 to 9 hours, while measuring a range of neurobehavioral performance functions each day (Belenky et al., 2003; Van Dongen et al., 2003a). Cumulative deficits in psychomotor speed and vigilance lapses were observed when sleep was below 7 hours per day, in accordance with the current definition of insufficient sleep. Importantly, the rate of performance deterioration was inversely related to the daily dose of sleep (i.e., the less sleep obtained per day, the more rapidly these deficits increased). Moreover, dividing the sleep into two periods each day (e.g., nocturnal sleep and daytime nap) did not have adverse effects on performance until the total time for sleep each day was below 7 hours (Mollicone et al., 2007, 2008), except that deficits when there was a nighttime circadian pressure for sleep did reveal performance deficits from not having a continuous sleep time (Mollicone et al., 2010).

There is some evidence of a (partially determined) genetically based difference among individuals as to sleep need, resistance to sleep deprivation, and circadian phase preference for sleep (see, for example, Goel et al., 2010; He et al., 2009; Pellegrino et al., 2014). However, the linkage between differences in phenotypic sleep need and vulnerability to sleep loss are not yet well enough understood neurobiologically to permit reliable prediction of behavioral risk. At this point, moreover, it is not possible to make use of genetic information to set more individualized hours-of-service (HOS) regulations. Therefore, this report makes no further mention of this issue.

WORK-RELATED INSUFFICIENT SLEEP IN COMMERCIAL MOTOR VEHICLE DRIVERS

Chronic insufficient sleep due to sleeping less than 7 hours per day results in a “sleep debt” (Van Dongen et al., 2003b) that gradually accrues over time, resulting in a greater tendency or propensity to fall asleep unintentionally (McKnight-Eily et al., 2009; Punjabi et al., 2003), and is associated with more reports of greater drowsiness when driving (Abe et al., 2012; Scott et al., 2007). Sleep debt can be reduced only by extending sleep time for recovery. Stimulants, such as caffeine, nicotine, and other drugs used to promote wakefulness, can increase arousal and transiently improve performance in fatigued individuals, but their effects are limited, and those who take them do not recover their sleep debt because there are no chemical substitutes for sleep (Bonnet, 2005; Spaeth et al., 2014).

Individuals who obtain less than 7 hours of sleep per day because of work and other activities typically sleep longer on nonworkdays to recover the sleep debt (Basner et al., 2014). This appears to be the case for CMV drivers as well. In the past 10 years, three naturalistic studies of CMV drivers have used wrist actigraphy devices to record objectively drivers' sleep times and sleep durations per 24 hours, on duty and nonduty days (Dinges et al., 2005b; Hanowski et al., 2007; Van Dongen and Mollicone, 2014). These studies revealed that the amount of sleep obtained by the drivers on workdays averaged 5.0 to 6.2 hours per 24 hours, while their sleep on off-duty days averaged 6.5 to 8.9 hours per day. In all three studies, the differences in sleep time on work- versus nonworkdays were statistically significant, and in all three, drivers' workday sleep durations averaged below 6.3 hours per day, an amount of daily sleep time considered insufficient for health (Watson et al., 2015a, 2015b). Thus, these studies suggest that on average, CMV drivers accumulate some degree of sleep debt due to insufficient sleep on workdays, and that they attempt to reduce that debt by sleeping 1-2 hours longer on off-duty days—a behavioral pattern of many adult Americans (Basner et al., 2014). Laboratory evidence indicates that a compensatory increase in sleep duration (i.e., being able to sustain sleep for a longer period of time) following 5 nights of sleep restriction occurs on 1 or more recovery nights in response to insufficient sleep (Banks et al., 2010). It is not known whether repeatedly cycling between 5-7 days of reduced sleep time per day and 1-2 days of extended (recovery) sleep has consequences for health and safety. Even if this pattern makes it possible for drivers to recover from workweek sleep debt, inadequate sleep during a workweek carries risks, as described below.

Fatigue from Inadequate Sleep

Studies have repeatedly shown that fatigue most often occurs as a result of the physiological consequences of inadequate sleep, prolonged wakefulness, and being awake at a circadian time when the brain is programmed to sleep. (The circadian nadir in alertness typically occurs between 11 PM and 7 AM, but can vary among individuals by a few hours, depending on habitual sleep timing.) These factors can co-occur to amplify the adverse effects of fatigue on performance and behavior.

Inadequate sleep also can result from lack of treatment of common medical conditions such as insomnia and sleep apnea, both of which have a high prevalence in the general population, although their frequencies in CMV drivers are not well established. Chronic insomnia is defined as the subjective perception of difficulty with sleep initiation, duration, consolidation, or quality that occurs despite adequate opportunity for sleep, and that results in some form of daytime impairment (Schutte-Rodin et al., 2008). Insomnia has been linked to lost work time and workplace accidents (Kucharczyk et al., 2012; Leger and Baron, 2010). If untreated, sleep apnea can result in excessive daytime sleepiness and risks to safety, especially when one is driving.

Sleep loss has a wide range of adverse effects on cognitive domains and neurobehavioral functions. These effects include (1) unstable attention, evident in both errors of omission (i.e., failure to respond in a timely manner to a stimulus) and errors of commission (i.e., responses when no stimulus or the wrong stimulus is present), as well as increased decrements in vigilance; (2) slowing of cognitive and psychomotor response times; (3) decline of both short-term and working memory performance; (4) reduced learning (acquisition) of cognitive tasks; (5) deterioration of performance in tasks requiring divergent thinking; (6) perseveration with ineffective solutions; (7) performance deterioration as task duration increases; and (8) growing neglect of activities judged to be nonessential (Goel et al., 2009).

Sleep Dose-Response Studies

Results of controlled laboratory experiments indicate that the effects of chronic sleep restriction are especially apparent when sleep is restricted to less than 7 hours a night (Belenky et al., 2003; Dinges et al., 1997; Van Dongen et al., 2003a). In these studies, performance deficits increased steadily across consecutive days of sleep restricted to less than 7 hours, and the less sleep was obtained per night, the more rapidly the performance deficits increased across days of sleep restriction. Within 5 to 6 days of sleep restricted to less than 7 hours, decrements in behavioral alertness increased to levels equivalent to having had no sleep at all for 24 to 48 hours (Van Dongen et al., 2003a, 2003b).

Sleep Propensity

In studies of sleep duration relative to daytime sleep propensity (i.e., the physiological tendency to fall asleep) and drowsy driving, adults reporting sleep durations of 6.75 to 7.5 hours and of less than 6.75 hours had a 27-percent and 73-percent increase, respectively, in the risk of sleep onset during the sleep propensity test compared with adults reporting more than 7.5 hours of sleep (McKnight-Eily et al., 2009; Punjabi et al., 2003). Studies also have shown that motor vehicle crash risk increases when self-reported sleep duration is less than 6 hours per day (Abe et al., 2012; Scott et al., 2007). A recent cross-sectional survey of drivers found an association between self-reported sleep duration of less than 7 hours per day and at least one self-reported incident of falling asleep while driving during the prior year.

Self-Reported Measures of Sleepiness and Fatigue

Dose-response studies on the adverse effects of sleep restriction on attention and performance have shown that self-reported sleepiness or fatigue does not continue to increase with chronic sleep restriction, but rather achieves a maximum at levels that do not reflect actual performance risks. Therefore self-reported sleepiness and fatigue may not reliably reflect increasing performance risks.

Sleep Inertia

Sleep inertia refers to a performance deficit that occurs upon awakening from sleep and that involves grogginess and a tendency to fall back asleep (Dinges, 1990). Although extended wakefulness without sleep can increase performance lapses, and sleep can reduce these effects, performance can be degraded by sleep inertia during the period immediately after awakening from sleep (for up to 2 hours, depending on the severity of prior sleep deprivation) (Dinges, 1990; Jewett et al., 1999). The more severe cognitive deficits of sleep inertia can be blocked by ingesting caffeine (Hayashi et al., 2003; Van Dongen et al., 2001).

Alcohol as a Contributing Factor

The fact that alcohol can impact the degree to which insufficient sleep affects a driver is becoming increasingly better understood. Alcohol appears to heighten the degree to which insufficient sleep impairs performance. (For further information, see Akerstedt et al., 2008.)

METHODS FOR REDUCING FATIGUE

Fatigue management in operational environments in general and in transportation modalities in particular is a major priority in many parts of the world, and the focus of a growing amount of human factors research (Abe et al., 2015). Countermeasures for fatigue are a special area of focus in fatigue management.[2]

---

[2] For an overview of fatigue countermeasures, see Williamson and Friswell (2013).

Chemical Countermeasures

The stimulant most commonly used as a fatigue countermeasure by many people, including CMV drivers, is caffeine, which acts pharmacologically to enhance alertness. Research has shown that caffeine has alerting effects and increases performance levels in the short term, especially in those who do not consume high doses on a regular basis (Nehlig, 1999; Neri et al., 1995). Caffeine affects the nervous system within 15-20 minutes, and its alerting effects can last for 4 to 5 hours, depending on the biological rate at which an individual clears it. However, research on the extent to which caffeine can maintain performance as sleep deprivation continues over days has revealed that caffeine can be ineffective for maintaining alert levels of performance as time awake extends past 16 hours and in individuals with high sleep debt (Spaeth et al., 2014).

Although CMV drivers may consume caffeinated foods and beverages, as well as ingest nicotine, then, it is important to reiterate that there is no biological substitute for sleep (Bonnet et al., 2005). The only way to recover effectively from sleep loss is to obtain adequate recovery sleep through a prolonged daily sleep period and the strategic use of naps (see below). Extensive evidence shows that sleep is the only reliable, natural, effective countermeasure for mitigating the neurobehavioral deficits due to sleep loss (Banks et al., 2010; Dinges et al., 1987; Mollicone et al., 2007, 2008; Rogers et al., 2003; Van Dongen et al., 2001).

Napping

A great deal of research conducted in both laboratory and operational settings has found that naps are an effective way to restore alertness and counter sleepiness when the time available for sleep is limited (see, e.g., Bonnet, 1991; Brooks and Lack, 2006; Costa, 1997; Driskel and Mullen, 2005; Lavie, 1986; Rosekind et al., 1994; Schweitzer et al., 1992; Tilley et al., 1982). Naps taken both prior to (i.e., "prophylactic" naps) and during (i.e., "power" naps) work periods have been found to improve alertness and performance relative to not napping (Bonnet, 1991; Dinges et al., 1987; Schweitzer et al., 1992). Moreover, a meta-analysis of 12 studies found that naps led to performance improvements that were directly proportional to nap duration (Driskell and Mullen, 2005). Other studies also have found a nap dose-response benefit (Bonnet, 1991; Brooks and Lack, 2006).

To maximize the beneficial effects of naps on performance and physiological alertness, aspects of nap timing must be considered. Daytime naps are typically difficult to initiate and maintain because this is the time at which the circadian clock is programming daytime alertness (Costa, 1997; Lavie, 1986; Tilley et al., 1982). During the afternoon "siesta" period, however, it is typically easier to initiate sleep in preparation for an upcoming night shift. Naps taken during the nighttime when the circadian clock is programming sleepiness are easier to maintain and show the most beneficial effects (Dinges, 1986). For example, a 1-hour nap at 4:30 AM was found to be more beneficial for next-day performance than a 1-hour nap at 9 PM (Gillberg, 1984). It is important to keep in mind that when awakening from naps taken during the circadian low point, individuals may experience a significant period of sleep inertia, or a feeling of grogginess, upon awakening, accompanied by decrements in performance.

Activity Breaks

People generally do not think of an activity break as an effective way to combat fatigue. However, research has shown that short breaks, especially when they include some physical activity, increase alertness by reducing the monotony of a task. The beneficial effects of breaks are due in part to postural changes that occur when one gets up and walks around (Caldwell et al., 2003; Dijkman et al., 1997; Heslegrave and Angus, 1985; Matsumoto et al., 2002). Studies have shown that short-term physiological and subjective alertness benefits lasting up to 15-30 minutes can be associated with activity breaks. However, performance on a test that is very sensitive to sleep loss (i.e., a psychomotor vigilance test) indicated no significant benefit of breaks (Neri et al., 2002). The beneficial subjective and physiological effects of the breaks were most pronounced around the low point in body temperature (i.e., the window of circadian low [Neri et al., 2002]).

Exercise

As suggested above, periods of light-to-moderate exercise have been shown to increase physiological arousal and help promote alertness (Buxton et al., 2003; Eastman et al., 1995; Horne and Foster, 1995; LeDuc et al., 2000; Stepanski and Wyatt, 2003). The beneficial effects of moderate exercise on subjective sleepiness have been shown to last up to 30 minutes (LeDuc et al., 2000), but improvements in objective performance have not been observed.

CONCLUSION

The extent to which CMV drivers utilize countermeasures for fatigue during their routine work schedules is unknown. The few studies that have monitored their sleep times objectively have found a considerable difference in their sleep on workdays and on days off duty. On workdays, they appear to obtain on average 6-6.5 hours per day, at the low end of what is considered sufficient sleep to maintain daily alertness and health. The additional sleep they obtain on days off duty (7-8.5 hours) suggests they are compensating for a cumulative sleep debt from the prior work week. The extent to which CMV drivers use naps, caffeine, nicotine, rest breaks, and exercise to mitigate their fatigue when working is unknown. Also unknown is whether their sleeping arrangements during work periods promote healthy physiological sleep continuity and duration, which are essential for optimizing waking alertness. Research is needed that can provide realistic estimates of the extent to which drivers are utilizing the above countermeasures, and the extent to which their use of those countermeasure facilitates alertness and performance.

4

Hours-of-Service Regulations

An important way in which the Federal Motor Carrier Safety Administration (FMCSA) tries to reduce crashes, injuries, and fatalities involving large trucks and buses is by issuing and enforcing hours-of-service (HOS) regulations for property- and passenger-carrying commercial motor vehicle (CMV) drivers in the United States. HOS regulations specify the maximum number of hours that truck and bus drivers are allowed to work in a 24-hour day or a 7-day week. The hope is that limiting the number of hours that CMV drivers are allowed to be on duty gives them the time to acquire sufficient rest, and by doing so reduces fatigue-related crashes. The regulations serve two purposes: (1) they limit time on task, and (2) they allow for periods of rest and sleep. This chapter provides a brief description of HOS regulations in the United States. It also describes HOS regulations in Canada and Mexico, since trucks and buses from these countries routinely enter the United States and vice versa, and HOS regulations in these countries can therefore be compared with those in the United States.

BACKGROUND

The first HOS regulations for interstate commercial drivers were issued in 1938 by the Interstate Commerce Commission (ICC), which was then the relevant regulatory agency. The ICC issued the regulations because it understood that the industry's operations generated safety concerns among the public, and the existing regulatory framework was falling short in addressing those concerns. The ICC subsequently modified its HOS regulations in 1939, 1962, and 1963, after which the regulations remained unchanged until 2003. In 1995, the ICC was abolished, and in the same act, Congress directed the Federal Highway Administration (FHWA) to establish new regulations incorporating the latest scientific knowledge about human fatigue and alertness.[1] The 1999 Motor Carrier Safety Improvement Act led to the creation of FMCSA and established safety as the new agency's highest priority. In 2003, FMCSA published new regulations that reflected the congressional directive, which were then revised in 2011. The regulations specify maximum on-duty time, minimum off-duty time, and maximum total driving hours. They also outline the manner in which a driver can utilize off-duty time to incorporate rest, which includes breaks from work and time for sleep.

DEFINITIONS

Duty Cycle

Duty cycle, in this context, refers to a day-like period of maximum on-duty time followed by minimum off-duty time, which may or may not sum to 24 hours. Thus, for example, if the maximum on-duty time were 12 hours and the minimum off-duty time were 10 hours, a "day" could restart after 22 hours, and a driver could do this until weekly limits had been reached. When the duty cycle differs from 24 hours, there is clearly some misalignment with the driver's circadian clock. A duty cycle differed from 24 hours until recently, when, to provide for circadian alignment, FMCSA set it to 24 hours. The term is sometimes also applied to limitations on the number of hours that can be spent driving in a 7- or 8-day period.

On-Duty Time

On-duty time refers to the time during which the driver is working, which encompasses driving, loading, and unloading. The HOS regulations set limits on on-duty and driving time:

- **On-duty limit** refers to the maximum amount of total duty time in a 24-hour duty cycle and a 7-day period that is allowed.
- **Driving limit** is the maximum amount of total driving time between two off-duty periods that is allowed.

---

[1] Interstate Commerce Commission Termination Act. See http://www.gpo.gov/fdsys/pkg/PLAW-104publ88/html/PLAW-104publ88.htm [March 2016].

Off-Duty Time

Off-duty time refers to the time during which the driver is not performing any work related to his or her job:

- **Off-duty limit** refers to the minimum number of hours that a driver must be off duty in a 24-hour duty cycle given that the driver is on duty for the maximum number of hours allowed by the HOS regulations.
- **Rest breaks** consist of the off-duty time that a driver utilizes for purposes of resting. In the United States, a rest break is a 30-minute period of time.
- **Sleeper berth provision:** If a truck has a sleeper berth (which should meet safety requirements), a driver can use the sleeper berth to rest during the off-duty period. The sleeper berth provision under the HOS regulations allows drivers some flexibility in how they utilize their off-duty time.[2]

In addition, the HOS regulations include a **restart provision** that requires drivers who drive the maximum number of allowable hours to restart their weekly duty cycle after being off duty for a certain number of hours.

PREVIOUS AND CURRENT HOURS-OF-SERVICE REGULATIONS

Table 4-1 summarizes the HOS regulations enforced by the ICC until 1995 and then by FMCSA starting in 1999.

The ICC and FMCSA also placed limitations on the maximum number of hours a CMV driver is allowed to drive in a weekly duty cycle. The HOS regulations from 1938 to 1962 allowed drivers to drive for 60 hours in a 7-day period or 70 hours in an 8-day period. The 2003 HOS regulations increased the maximum driving hours to 77 hours in a 7-day period or 88 hours in an 8-day period, while the 2011 regulations reverted back to 60 or 70 hours in a 7- or 8-day period, respectively.

Table 4-1 reveals that there has been little change in the distribution of off-duty minimums and on-duty maximums in almost 80 years. The regulations of 1962, 1963, and 2003 required drivers to work in a cycle of work (driving time) and rest (off duty) that was less than 24 hours (i.e., an 18-hour rotation in 1962-1963 and a 21-hour rotation in 2003). Such duty cycles ignored the nondriving tasks performed by drivers while on duty,

---

[2] For more details on the sleeper berth provision, see https://www.fmcsa.dot.gov/regulations/title49/section/393.76 [March 2016].

TABLE 4-1 Hours-of-Service Regulations

Enforcement Year	Minimum Duty Cycle (hours)	On-Duty Maximum (hours)	Driving Maximum (hours)	Off-Duty Minimum (hours)	Sleeper Berth Provision	Restart Provision
1938	24	15	12	9	None	None
1939	24	Rescinded	10	8	Can split 8 hours of off-duty time into two periods.	None
1962	18	15	10	8	Can split 8 hours of off-duty time into two periods provided neither is less than 2 hours long.	None
1963	18	15	10	8	Can split 8 hours of off-duty time into two periods of 4 hours each.	None
2003 (applicable only to truck drivers)[a]	21	14	11	10	Can split 10 hours of off-duty time into periods of at least 2 hours.	Restart weekly on-duty hours after 34 consecutive hours off duty; on-duty time limit is 60 hours in 7 days or 70 hours in 8 days.
2011 (applicable only to truck drivers)[b]	24	14	11	10	8 consecutive hours in sleeper berth, plus a separate 2 consecutive hours either in sleeper berth, off duty, or a combination of the two.	The 34 hours off duty must include two periods from 1 AM to 5 AM; use of restart provision limited to once a week.

2011 (applicable only to bus drivers)	24	15	10	Not applicable	Same as above	Not applicable
2013 (applicable only to truck drivers)[c]	24	14	11	10	Same as above	Suspended

---

[a] 49 CFR Parts 385, 386, 390.

[b] See http://www.fmcsa.dot.gov/regulations/rulemaking/2011-32696 ; https://www.gpo.gov/fdsys/pkg/FR-2011-12-27/html/2011-32696.htm [March 2016].

[c] See http://www.fmcsa.dot.gov/regulations/hours-service/summary-hours-service-regulations [March 2016].

SOURCE: Adapted from Goel and Vidal (2014, Table 1).

and as noted above, they were not aligned with the circadian (24-hour) biological cycle regulating sleep and alertness. Between 1963 and 2003, FHWA issued two notices of proposed rulemaking, one in 1978 and the other in 1992, to change the HOS regulations. However, neither of those proposals was finalized.

FMCSA's 2003 rulemaking took into account the research done on the human need for recovery sleep by introducing the restart provision, which, as noted above, allowed drivers to reset their weekly duty cycle. These two changes in the 2003 HOS rules, which lengthened the allowable number of hours per week to greater than 60 and at the same time instituted a restart provision, are somewhat at cross-purposes, since one change reduces the potential off-duty time for workdays, while the other maintains it for nonworkdays. FMCSA's 2011 rulemaking changed the duty cycle to 24 hours to be consistent with the circadian cycle. It also modified the 34-hour restart period to include two consecutive nighttime periods (encompassing the time interval from 1 AM to 5 AM each night) that a driver was off duty between two weekly duty cycles. This was done to increase driver sleep time during circadian and environmental night. The 2011 rulemaking also rescinded the flexibility to split sleep under the sleeper birth provision so as to discourage drivers from experiencing irregular and shorter periods of sleep. FMCSA indicated that the new rules "moved drivers toward a work-rest schedule that more closely matched the natural circadian cycle of 24 hours and gave drivers the opportunity to obtain the 7 to 8 hours of uninterrupted sleep per day that most adults need," which would be concordant with the discussion of driver fatigue and human physiology in Chapters 1 and 3.

It is obvious from Table 4-1 that truck and bus drivers need to manage their time and track their activities to avoid violating the HOS regulations. The regulations are promulgated for safety purposes and are applied uniformly to both the truck and bus industries. However, the two industries are highly segmented and heterogeneous in terms of the demands placed on their drivers (see Chapter 2), so the impact of the regulations depends on the type of trucking or busing in which a driver is engaged. Therefore, to understand the linkages among driving times, driver fatigue, and crash risk, both current HOS regulations and drivers' typical driving schedules are among the many factors that need to be accounted for. Part III of this report describes the research that has been conducted to investigate those linkages.

HOURS-OF-SERVICE REGULATIONS IN CANADA AND MEXICO

Regulations for work and rest hours for truck and bus drivers have also been promulgated by Canada and Mexico. The combined trade relationship among the United States, Canada, and Mexico amounts to $1.4 trillion annually, making trucking a major contributor to commerce among the three countries. And just as the movement of goods is facilitated by trucks, the mobility of persons is facilitated by bus networks that enable travel for people from the three countries.[3] The regulations for Canada and Mexico, as well as Australia and the European Union, are summarized in Table 4-2.

TABLE 4-2 Hours-of-Service Regulations in Other Countries

Country	Minimum Duty Cycle (hours)	On-Duty Maximum (hours)	Driving Maximum (hours)	Off-Duty Minimum (hours)
Canada[a]	24	14	13	8[b]
Mexico	None	Up to 3 hours of daily overtime (up to 9 hours/week)	8 hours (6 AM-8 PM) 7 hours (8 PM-6 AM) 7.5 hours (if two shifts in 24 hours)	None
European Union[c]	24	14.25	10	9
Australia[d]	24	12	12	7

[a] Drivers operate on "cycles." Cycle 1 is 70 hours in a 7-day period. Cycle 2 is 120 hours in a 14-day period. A driver who uses Cycle 1 must take off 36 hours at the end of the cycle before he or she can restart the cycle again. A driver following Cycle 2 should be off duty 72 hours before starting on a cycle.
[b] An additional 2-hour period of rest that must not be taken in less than 30-minute blocks.
[c] The weekly limit in the European Union is 56 hours.
[d] The weekly limit in Australia is 72 hours.

SOURCE: Adapted from Goel and Vidal (2014, Table 1).

---

[3] Annual data on Border Crossing/Entry of Trucks and Buses from Canada and Mexico, U.S. Department of Transportation, Research and Innovative Technology Administration, Bureau of Transportation Statistics, based on data from the Department of Homeland Security, U.S. Customs and Border Protection, Office of Field Operations.

PART II

CURRENT RESEARCH DATA AND METHODS

5

Data Sources

A major objective of transportation agencies is to improve road safety. Accordingly, data on driver behavior and vehicle performance are a vital tool in formulating policies and designing systems to achieve that objective. The purpose of this chapter is to provide an understanding of the wide range of data sources that are potentially available to researchers for identifying factors that can reduce risk and enhance safety in the transportation network. For each source, the strengths and limitations of the data are considered, especially with respect to answering key research questions about fatigue among commercial motor vehicle (CMV) drivers. The sources reviewed include crash databases, naturalistic driving studies, and driving simulator studies. Both publicly available and private resources are described; the latter would require interaction with the data holders to be used for research purposes. Also discussed are promising data sources in development that have the potential to better inform the policy-making process of transportation agencies.

OUTCOMES AND PREDICTORS

Research analyzing the relationships among hours driven, driver fatigue, and highway safety has relied on the paradigm of the Haddon Matrix (the most commonly used paradigm in the injury prevention field) to recognize factors that lead to different safety outcomes (Haddon, 1972). Safety outcomes include crashes and noncrashes, while predictors can be broadly categorized into driver characteristics, vehicle characteristics, carrier characteristics, and environmental factors. The data sources described below collect information on various aspects of safety outcomes and many of their predictors, but there is no single repository of these databases. Also, it is unknown whether there are identifiers in these databases that can facilitate linking across different files. This linking is essential since one must have information available for each unit of analysis, and these units are travel segments for a given vehicle. For each segment, therefore, one must have the relevant information on the predictors and the outcomes, often crashes, so that the causality of individual factors can be analyzed. Having marginal information on a potential causal factor does little good since one must know how it varies according to the other potential causal factors.

PUBLICLY AVAILABLE COMMERCIAL MOTOR VEHICLE CRASH DATABASES

Databases containing information on CMV crashes are maintained primarily by the National Highway Traffic Safety Administration (NHTSA) and the Federal Motor Carrier Safety Administration (FMCSA). They are used to create national- and state-level estimates of CMV crash rates disaggregated by various characteristics of interest, such as fatigue. A key starting point for most of these databases is police accident reports. The following subsections describe these databases.

Fatality Analysis Reporting System (FARS)

FARS is a census file of motor vehicles involved in fatal traffic crashes in the United States. FARS data are compiled and maintained by the National Center for Statistics and Analysis within NHTSA. Data for the FARS file are collected by analysts within each state from police accident reports, death certificates, medical examiner reports, hospital reports, emergency medical services reports, state vehicle registration files, state driver licensing files, state highway department data, and other information from any follow-up investigations. Analysts use the information from these sources to code 100 data elements on certain events leading up to a crash, including crash characteristics, environmental conditions, driver distractions, circumstances obscuring the driver's vision, and description of persons (e.g., age, gender, restraint use, injury severity) and vehicles (e.g., type, make/model, model year, cargo body for trucks) involved in the crash. Driver fatigue is coded, if specifically identified on the crash report, and is included in the data as part of a field that records driver impairments.

Trucks Involved in Fatal Accidents (TIFA) and Buses Involved in Fatal Accidents (BIFA)

TIFA was a census of medium and heavy trucks involved in fatal crashes, while BIFA was a census of buses involved in such crashes. TIFA and BIFA data were collected by the University of Michigan Transportation Research Institute (UMTRI) with support from FMCSA. Both of these data sets were based on the FARS file, supplemented by data collected by UMTRI researchers. The researchers used telephone interviews with drivers, police officers, dispatchers, emergency personnel, witnesses, and others with knowledge of the trucks or buses involved in the fatal crashes. Data collected included a detailed description of the truck or bus, the carrier and carrier operations, driver hours and compensation, and details about the initiating events of the crashes. This information was significantly more detailed than the general-purpose data collected by FARS. TIFA included data on hours driven prior to a crash and the intended trip length (which could serve as a surrogate for intended time on task). Both the TIFA and BIFA surveys were discontinued after 2010.

National Automotive Sampling System (NASS)

NASS has two components—the General Estimates System (GES) and the Crashworthiness Data System (CDS). GES is a general-purpose crash data file of motor vehicle crashes in the United States. CDS is focused on light vehicles and is not discussed further here. Unlike FARS, CDS includes all crash severities; unlike TIFA or BIFA, it includes all motor vehicle types.

The GES database is based on a hierarchical stratified sample of police-reported crashes that involved at least one motor vehicle traveling on a roadway with resulting property damage, injury, or death. GES data are used to provide national-level estimates of a comprehensive set of descriptors on crashes. GES data collectors make weekly visits to approximately 400 police jurisdictions in 60 Primary Sampling Units across the United States. Approximately 50,000 police accident reports are sampled each year. Data items in GES are coded entirely from those reports. Given the nature of the sampling structure, standard errors can be relatively large for small subsets of the data, although a significant strength is the consistency and comprehensiveness of the data.

Motor Carrier Management Information System (MCMIS)

FMCSA maintains the MCMIS crash file, a census of all trucks and buses involved in a crash that included a fatality, an injury transported for immediate medical attention, or at least one vehicle towed because of disabling damage. Qualifying vehicles include trucks with a gross vehicle weight rating (GVWR) greater than 10,000 lb and buses designed to transport more than eight people, including the driver. States are required to extract and submit a standard set of data elements for each qualifying vehicle involved in a traffic crash that meets the crash severity threshold. The MCMIS crash file, carrier file (registration data on all qualifying motor carriers), and inspection file (data from all inspections of motor carrier vehicles and drivers) make up the MCMIS. The carrier file includes for each carrier, along with other information, the estimated vehicle-miles traveled (VMT); the number of trucks operated, disaggregated by ownership (owned, leased, trip leased); power unit type (tractor, straight truck); counts of certain trailer types; and counts of drivers, employed and leased. This information is provided by the carriers. The carriers are required to update the information every 2 years; carrier information also is updated as part of safety audits. These three different files can be linked using the carriers' U.S. Department of Transportation (DOT) numbers.

FMCSA uses the MCMIS to monitor safety levels and identify unsafe carriers for interventions. The data are used primarily to evaluate carriers in terms of their crash performance; safety performance; and compliance with regulations, including hours-of-service (HOS) regulations. The data are less useful for detailed scientific evaluation of specific safety questions. Data elements on the individual driver and the truck/bus involved in a crash are limited; therefore, controlling for confounding factors (such as operational conditions) often is not possible in an empirical analysis. The exposure data in the carrier file are aggregated at the carrier level, so one cannot use these data to compute crash rates by environmental, road, or driver conditions. A researcher can theoretically obtain information on the driver's condition (especially fatigue) from the police report, but doing so often is not possible in practice because MCMIS and state crash files typically include no common identifier. Carriers can challenge data on specific crashes through the DataQ Program, run by FMCSA, which allows a review of the data issued by FMCSA. Through this process, the system automatically forwards a request for data review to the appropriate office for resolution and collects updates and responses for current requests. Carrier data are checked during audits and some other enforcement activities, but only a small number of carriers are subject to these processes. In addition, FMCSA runs the Safety Data Improvement Program to help states improve their reporting. However, no systematic program is currently in place with which to evaluate the comprehensiveness and consistency of the data.

Given the diversity of the trucking and bus industries, it would be useful for FMCSA to support research that could provide a better under-

TABLE 5-1 Coverage of Crash Severity by Commercial Motor Vehicle Crash Databases

Crash Severity	FARS/TIFA	GES	MCMIS
Fatal	Yes	Yes	Yes
Injury	No	Yes	If injured person(s) transported for medical attention or if any towed/disabled vehicle
Property damage only	No	Yes	If any towed/disabled vehicle

NOTE: FARS = Fatality Analysis Reporting System; GES = General Estimates System; MCMIS = Motor Carrier Management Information System; TIFA = Trucks Involved in Fatal Accidents.

standing of which information is and is not included in the MCMIS. One way of doing so would be to compare the MCMIS with the sampling frame used by the Bureau of Labor Statistics' Survey of Occupational Employment Statistics.

The coverage of crash severity in FARS/TIFA, GES, and the MCMIS is summarized in Table 5-1.

Limitations of CMV Crash Databases

The data in the FARS, TIFA, GES, and MCMIS crash databases all begin with police crash reports. When using these data, it is important to bear in mind that police officers' primary function is protecting the public and enforcing the law, not collecting data for scientific studies. Thus driver fatigue at the time of a crash is identified either by driver admission, witness observation, or inference based on vehicle crash characteristics and driver work-rest schedule prior to the crash. Currently, there are no objective tests for fatigue that can feasibly be administered when crashes occur, analogous to tests for drugs or alcohol. Because the identification of a fatigued driver in these databases depends on the reporting police officer's identifying and recording fatigue,[1] CMV driver fatigue likely is underreported in existing crash data. An independent assessment of the reporting of fatigue in GES suggests that 1.4 to 3.1 times more crashes involve fatigue than are reported in this database (Knipling and Wang, 1995). Moreover, crash data in these databases include no information

---

[1] The Federal Aviation Administration has developed a Fatigue Risk Assessment Tool, which is freely available on the agency's website and is to be used by aviation maintenance workers to capture sleep- and fatigue-related inputs relevant to a maintenance incident. This tool ensures that all incident investigations follow a uniform definition of fatigue. A similar tool might be useful for the CMV industry.

about recent hours of sleep and quality of sleep obtained by the driver, and with the termination of the TIFA and BIFA projects, no data are available on hours driving since the last break.

Even with their limitations, these crash databases have their utility. First, they directly address the safety issue—that is, crashes. Second, while it is unlikely that all fatigue-related crashes are identified, these databases provide the most direct measure of the effects of driver fatigue on safety relative to hours of service.

LARGE TRUCK CRASH CAUSATION STUDY (LTCCS) DATABASE

The LTCCS was a collaborative project of FMCSA and NHTSA, aimed at collecting detailed information on crash events and associated causal factors. For this study, 963 large-truck crashes were selected using a clustered, stratified sampling procedure. The sampled crashes involved a total of 1,123 trucks and 932 other vehicles and took place between 2001 and 2003. Sampled crashes involved a least one truck with a GVWR of more than 10,000 lb and at least one fatality, incapacitating injury, or non-incapacitating but evident injury. Cases were sampled from 24 sites across the country to produce a nationally representative crash file. It should be noted that, because of the relatively small number of crashes and vehicles included in this study, standard errors for population estimates are large. Each case was investigated by a trained researcher, many experienced in the NASS CDS. In addition, a Commercial Vehicle Safety Alliance (CVSA)-trained truck inspector performed a CVSA Level 1 inspection[2] of the trucks involved in most of the crashes for the study. The researchers conducted in-depth investigation for each of the crashes included in the study, encompassing scene diagrams; photographs; and extensive information on the vehicles, environment, and drivers.

The crash data collection was structured around precrash maneuvers, the critical event itself, the critical reason for the critical event, and associated factors. For details, see Blower and Campbell (2002), Findley et al. (2000), and Hedland and Blower (2006). These factors (especially the definition of critical event and critical reason, along with the nature of the associated factors) are sometimes misunderstood. The "critical event" was defined as the event that made the crash or collision unavoidable. The "critical reason" was the reason for the critical event, that is, the last failure or error that precipitated the crash. For example, in a crash in which a light vehicle turned across the path of a truck at an intersection and the truck could not evade it, the critical event would be the light vehicle's

---

[2] For information on the North American Standard Inspection Program, see http://www.cvsa.org/programs/nas.php [March 2016].

turning. The critical reason would be assigned to the light vehicle and would be attributed to whatever failure was responsible—for example, the light vehicle driver did not see the truck or misjudged the gap.

Fatigue was captured both as a critical reason, when determined to be the reason for the precipitating event, and as an "associated factor," when present regardless of its contribution to the crash. In addition, the researchers coded data on hours spent driving prior to the crash and hours on duty, typically obtained from driver log books. They attempted to validate such driver log books using any other evidence available, such as the time stamps on fuel and food receipts and tolls. They also tried to obtain data on the hours of last sleep prior to the crash and the driver's schedule for the 7 days prior to the crash. In some cases, the researchers were on the crash scene and could directly evaluate the driver's condition. The assessment of fatigue was based on a reasonable judgment of the totality of available evidence.

Even though the data are more than 10 years old, the LTCCS is still the most comprehensive and detailed truck crash investigation data set available. (For more information on this study, see National Highway Traffic Safety Administration and Federal Motor Carrier Safety Administration, 2006a, 2006b, 2012.) The LTCCS data set has significant advantages over other crash data sets. Overall, it is much richer than the others. Data were captured for a comprehensive list of factors that could have contributed to the crashes. Also, use of the critical event/critical reason paradigm supports a flexible analytic approach (see Hedlund and Blower, 2006). Further, the researchers' narratives provide a rich source of information. For purposes of the linkage between fatigue and crash risk, the researchers' assessment of fatigue was better than any available crash data based on police reports (such as FARS, GES, or state crash reports).

However, it should be recognized that some of the information related to fatigue is self-reported, particularly a driver's hours of sleep prior to the crash. Researchers had little reasonable opportunity to evaluate whether drivers actually had slept the hours claimed. In addition, since duty hours were collected from log books, even given that investigators used common techniques for verification, there was the potential for misreporting. Also, while the assessments of fatigue (i.e., driver sleep and alertness) were almost certainly better than those in other crash data sets, they were not as valid and reliable as the information that could be obtained using objective measures of sleep by wrist actigraphy and of alertness by performance on a psychomotor vigilance test. The primary disadvantage of the LTCCS data set is that the last crash in the data set occurred more than 10 years ago. Since then, many circumstances have changed, such as the HOS regulations, the ability to collect telematics information, the introduction of crash avoidance technologies, and distraction due to cell phone usage. Further, there are no associated exposure data, so it is not possible for analysts to control for many potentially confounding factors (a shortcoming that is common to all crash data).

DATA ON VEHICLE-MILES TRAVELED

Crashes are the product of risk times exposure. Risk is a function of characteristics of the driver, the carrier, the environment, and the vehicle. However, crashes are also a function of exposure—that is, time or distance driven. Thus, measures of exposure are needed to determine crash risk, but valid and useful exposure data are typically lacking. The main source of national exposure data is provided by the Federal Highway Administration (FHWA). Table VM-1 in *Highway Statistics* provides VMT statistics by vehicle and roadway type. Trucks are classified as a single unit or as a combination (pulling a trailer). The tabulations are produced at an aggregate level (at the national and state levels by type of vehicle and roadway); they currently are not broken down by carrier type or even by time of day. This deficiency in sources of exposure data is exacerbated by the fact that the Bureau of the Census discontinued the Vehicle Inventory and Use Survey (VIUS) in 2002. VIUS was a survey of truck owners that provided aggregate travel and use data over the course of a year. FHWA is considering filling this information gap by working with the Bureau of Transportation Statistics to carry out the necessary data gathering. As part of an effort now under way, FHWA is planning to produce statistics on VMT by hour of day, day of month, and month of year by major vehicle groups.[3]

It is important to understand that, with the exception of the LTCCS, the available crash data are collected as part of administrative systems that exist for purposes other than answering scientific questions. Crash data are collected to monitor overall levels of safety, to enforce the law, and to allocate public resources for reducing the toll of motor vehicle crashes. Similarly, the limited exposure data available, such as VMT, are collected to monitor overall trends in highway usage and to allocate highway funds to the states. These systems were not designed by safety scientists to answer scientific questions related to HOS, CMV driver fatigue, and safety.

---

[3] Presentation by Tianjia Tang from the Federal Highway Administration to Panel on Research Methodologies and Statistical Approaches to Understanding Driver Fatigue Factors in Motor Carrier Safety and Driver Health, September 3, 2014, Washington D.C.

RESEARCH- OR STUDY-BASED DATA SETS

Research- or study-based data sets include naturalistic driving studies (NDS) and driving simulator studies.

Naturalistic Driving Studies

Observational or naturalistic data collection is in situ, that is, "in the natural or original position or place." In NDS, data on driver behavior and performance are collected in the normal operating environment. For truck and bus NDS, the operational environmental reflects revenue-producing operations, and study participants operate their vehicle as part of their regular job duties. Study participants volunteer to have their vehicles instrumented while they go about whatever driving they would have performed otherwise. Typically, NDS are designed such that study drivers use an instrumented vehicle for an extended period (often several weeks or longer); data are collected continuously via the instruments until the engine is turned off. Instruments such as video cameras and other sensors collect data on the position and performance of the vehicle relative to the road and other vehicles in the vicinity, as well as driver behavior. The final data set provides an "instant replay" of the entire driving trip, including any incidents, allowing researchers to focus on event factors including driver behavior and crash precursors. NDS have been conducted with light vehicles (e.g., the 100-car study [see Dingus et al., 2006]), long-haul trucks (e.g., Hanowski et al., 2009), local/short-haul trucks (e.g., Hanowski et al., 2003), buses (ongoing), motorcycles (Williams and McLaughlin, 2013), and bicycles (Dozza et al., 2013).

Early NDS, such as the 100-car study (Dingus et al., 2006) and local/short-haul study (Hanowski et al., 2003), involved a relatively modest number of instrumented vehicles—100 light vehicles in the 100-car study and 4 trucks in the local/short-haul study. However, as instrumentation costs have decreased and the robustness, capacity, and ease of installation of data collection systems have improved, newer studies have included much larger numbers of vehicles. For example, the Transportation Research Board's Strategic Highway Research Program (SHRP) 2 involved the instrumentation of 3,353 vehicles and data collection from 3,546 drivers (McClafferty et al., 2015). In total, the SHRP 2 NDS resulted in approximately 50 million miles of continuous driving data.[4] Recent and ongoing NDS with trucks and buses have similarly increased in scope relative to early NDS. For example, an NDS involving commercial trucks

---

[4] SHRP 2 Products Chart. Available: http://onlinepubs.trb.org/onlinepubs/shrp2/SHRP-2ProductsChart.pdf [April 2016].

and buses funded by FMCSA, soon to be concluded, has been collecting continuous naturalistic data from 161 trucks (169 drivers) and 43 buses (68 drivers). These large NDS data sets have resulted in tremendous research opportunities and have led to the development of new online databases in which the NDS data can be accessed. The most recent of these databases that have come online include data from the SHRP 2 study;[5] at the InSight Data Access Website, some of the SHRP 2 data can be accessed for analyses.

Operational definitions of driving segments can vary across studies, and the trigger criteria used to identify them also can differ based on the research question of interest. In this way, NDS that use continuously recorded data offer the flexibility to customize search criteria to enable identification of events of interest to include in the analyses. With respect to the duration of interest, for example, analyses outlined in the 100-car study (Dingus et al., 2006) focused on a time frame of 6 seconds. By contrast, more recent analyses of light vehicle data focused on assessing driver behavior prior to an event trigger have reviewed video and other driving data 12 seconds prior to the event (Victor et al., 2014). Analyses of metrics associated with driver fatigue (e.g., PERCLOS, or the percentage of closure of the driver's eyelids over some time period) have used much longer time frames for video review; for example, Hanowski and colleagues (2013) reviewed 3 minutes of video prior to an event trigger to assess PERCLOS.

As defined for the 100-car study and the SHRP 2 study (Guo et al., 2010; Victor et al., 2015), a crash event is "any contact that the subject vehicle has with an object, either moving or fixed, at any speed in which kinetic energy is measurably transferred or dissipated,"[6] while a near-crash event is "any circumstance that requires a rapid, evasive maneuver by the subject vehicle, or any other vehicle, pedestrian, cyclist, or animal to avoid a crash."[7] Trigger criteria include a host of kinematic variables that act as filters by which a driving segment is selected from the video data for further review. Within a study, the objective is to select trigger criteria that reduce the chances of false alarms. Using different trigger criteria, various sets of events of interest can be captured.

Analysis of these large data sets has required new, and evolving, analytical approaches. Often, the data are characterized by events of interest (e.g., "crashes") and compared with baseline, or normative, driving.

---

[5] See https://insight.shrp2nds.us [March 2016].

[6] Examples of objects include other vehicles, roadside barriers, and objects on or off of the roadway, pedestrians, cyclists, or animals.

[7] A rapid, evasive maneuver is defined as steering, braking, accelerating, or any combination of control inputs that approaches the limits of the vehicle capabilities.

In this way, researchers can study the pre-event behaviors of the driver during the time of interest leading up to the event. In a similar way, the analysis can use nonevent data for comparison. In the SHRP 2 data set, 1,541 crashes were recorded, and although these severe outcomes can serve as one event type of interest to study, less severe events that may involve underlying factors of interest (e.g., driver distraction, speeding, rural roads) also can be studied by examining other safety events with outcomes less severe than crashes. For example, many NDS studies analyze incidents often referred to as "safety-critical events" (SCEs), which may include near-crashes and other driver errors (e.g., unintended lane deviations). These event types, in addition to matched nonevents that serve as a baseline for comparison, can be a focal point for video and data review. Further, as the flexibility of the above definition of triggers suggests, one can and should examine the robustness of any inference to the definition of the trigger one is using (which in turn defines the length of the period prior to an event on which one collects data).

The strengths of NDS as an approach for data collection are as follows:

- **Ecological validity:** Data are collected on driver performance and behavior, as well as vehicle behavior, including the factors that can jointly cause a traffic crash. NDS can provide a way of observing all of the relevant factors acting jointly as they might be when the vehicle is not instrumented. Assuming that the instrumentation has no impact on driver behavior—there is some evidence to suggest this is the case after only a few days of use of the instrumented truck—one gets to observe that behavior in a true operational environment. This NDS advantage provides very high ecological validity. There is no pressure for a driver to behave unnaturally since the experimenter is not present physically, and has given no instructions on how to drive, which roads to travel on, how to deal with excess traffic, or how to handle the environmental conditions encountered. In addition, in NDS, one can observe the results of typical driver pressures to, for example, make appointments; driver choice behavior (e.g., where to purchase fuel, whether to use caffeinated or tobacco products or eat or drink alcohol or other beverages shortly before or while driving whether to drive in adverse weather conditions); the time of day driving starts; the timing of waking and sleep (onset-offset times, duration) in advance of driving; and the like. All such decisions are up to the drivers under study, and drivers are allowed to behave as they would routinely.
- **Reduced errors in data collection:** Since the vehicles are instrumented, NDS provide precise, detailed information on driver

behavior and driving performance for occasions that both do and do not lead to crashes or SCEs. Review of video and other data allows for "instant replay" capabilities that make it possible to better understand the genesis of crashes and SCEs. Data are collected not only on crashes but also on events that under some circumstances are likely to be associated with a higher chance of having a collision. Such SCEs may include excessive lane changes, hard braking, tailgating, and speeding. Data also are collected on the behavior of other vehicles in the vicinity.

- **Case control and exposure:** One can easily carry out case-control studies using data from NDS. An important benefit of NDS data collection is that data exists for travel times when crashes or SCEs did not happen.[8] Since data are available for both crashes and SCEs, one can match a situation in which a crash or SCE occurred with an analogous situation in which one did not occur, and then examine the frequency of various possible causal factors to see whether it differs between the two situations. The data contain precrash and pre-SCE information on driver behaviors including the presence of fatigue or distraction and interactions with passengers, devices, and the vehicle.
- **Utility:** Finally, while NDS are expensive to conduct, they typically involve collecting data for months or longer, which allows for analyses investigating many research questions.

NDS also have limitations, although some of these can be addressed:

- **Events of interest:** In NDS the investigator has very little control over the driver, and therefore over the driving situations encountered. Thus one cannot study specific problems as intensively as is possible with simulators since those problems may not occur sufficiently often during the course of the study. This limitation is a standard by-product of high ecological validity. Also, since crashes are rare events, analysts tend to rely on SCEs. However, the validity of treating such events as near-crashes is uncertain. This is clearly less of a limitation with larger NDS such as SHRP 2, which included 701 crashes as of June 30, 2014 (Antin et al., 2015, Table 4.2). Also, relying on some SCEs may not be a serious problem in every situation, as analyses linking crashes to near-crashes have been conducted (see Guo et al. [2010] for an example). NDS also can be used to conduct field operational tests for technology evaluation (e.g., Hickman and Hanowski, 2011).

---

[8] SCEs can also provide information on how drivers successfully avoided crashes.

- **Feature extraction:** One of the reasons SCEs are defined as high-g-force events is that one can then focus attention on a small fraction of the data collected. The problem of feature extraction, if one suspects other causal factors in addition to these kinematic events, either requires going carefully over thousands of hours of video data capture or finding some other way to identify those moments. However, this may not be an issue for certain analyses that may look for events of interest other than SCEs. For example, if speeding over a set threshold (e.g., 80 mph) were of interest for analysis, search criteria could be set to identify all instances of high-speed events.
- **Causality:** NDS are observational studies and as such are subject to traditional limitations in their capability to address causality. In most cases, researchers do not have detailed information about driver state (fatigue, alcohol or drug use, and medical conditions that impair driving), which could introduce confounding factors.
- **Generalizability:** Because little information is available to identify differences in the characteristics of drivers of instrumented and noninstrumented vehicles, it may not be possible to determine whether such differences might confound any observed relationships. A related issue is that the subjects are volunteers, and it is reasonable to expect that those who engage in activities that affect their driving negatively would be less likely to participate. However, being a volunteer is an inherent component of almost all research involving human subjects, so the same limitation is inherent in all empirical studies. Therefore, extrapolating the results of NDS to the general trucking population is not entirely feasible, especially in the absence of baseline estimates, although NDS drivers can be compared demographically with other populations of drivers.
- **Sample size:** As noted previously, early NDS often had small sample sizes, although it is important to note that the same is true of many other empirical studies. However, the number of drivers in a small sample limits the variability one can observe in terms of fleet size, operation types, and vehicles. Again, this limitation is not unique to NDS; surveys and studies based on paper logs, for example, are restricted to certain segments of the CMV population. Achieving a study sample that is representative of the population is hampered by the fact that estimates of characteristics of all CMV drivers are not available. On the other hand, NDS provide an opportunity to study the same drivers over many months of driving.

In some cases, these limitations of NDS can be addressed through the use of additional data sets or modeling approaches. An example is the inability to assess driver sleep hours. Drivers in the Hanowski et al. (2007) study wore actigraphy devices that recorded their sleep quantity (timing and duration). These data were then merged with the driving data to better understand the impact of sleep on driver behavior and risk. Other data streams and data sets—such as data from geographic information systems, including data on weather and traffic density—can be combined with naturalistic driving data (see, e.g., Cannon and Sudweeks, 2011).

Driving Simulator Studies

Driving simulators are useful for research purposes as they allow a researcher to observe driver behavior under certain conditions while controlling for others. They are especially useful for the study of high-risk conditions, such as fatigue and other types of impairments, because performance errors do not result in injuries or fatalities. However, that fact also can limit the realism of the simulator experience. Simulators are used as well to test potential road and vehicle improvements and other technologies before they are deployed and obtain practical feedback from drivers. For example, simulators guided the design of advanced driver support systems by making it possible to estimate the marginal effect of their deployment on driver performance. Professional driving simulators also can serve as a training tool for individuals learning to drive trucks and buses. Depending on their fidelity to an actual truck or bus, they can provide approximate on-the-job training for a driving student and possibly reduce crash risk. Simulators are useful as well for investigating specific issues, such as whether fatigue is associated with unintended lane changes on curved highways, and for recognizing factors that need to be examined in broader studies. One can look at a variety of outcomes by repeating the experiment with a study participant, thereby controlling for individual differences. In addition, a researcher can investigate low-frequency/high-severity events such as a crash by repeating the same scenario without adding substantial costs. Not only are driving simulators useful for testing technologies but they also may have utility for testing various types of schedules to identify which is optimal for truck and bus drivers—although the lack of fidelity to the real world may limit the utility of this application. Finally, simulators give researchers the flexibility to gather additional information such as physiological data (e.g., heart rate), although more miniaturized portable physiological monitoring with data uplink is rapidly becoming possible, so this advantage is becoming less clear.

Simulator studies are limited in the sense that their results are difficult to generalize, and the estimate of an impact of a factor will shift upward or downward depending on how it interacts with other factors in the real world (that were fixed or controlled for in a simulator study). Finally, as the same experiment is repeated with a participant in a simulator study, the effect of learning becomes a confounding factor.

PROPRIETARY DATA

Proprietary data include data collected by the American Transportation Research Institute (ATRI) and by large truck carriers.

Data Collected by the American Transportation Research Institute

ATRI is a member of the American Trucking Associations (ATA) and is a not-for-profit research organization. ATA is organized as a federation of independent state motor carrier associations, councils, and committees, and ATRI was established as a separate organization to maintain its independence from ATA. In essence it is the research arm of ATA, with its own management (a separate board of directors consisting mainly of ATA members) and funding, which comes from the trucking industry. ATRI conducts studies in various topical areas relevant to the trucking industry and therefore collects data on drivers and motor carriers. It conducted surveys on fleet managers and drivers and also made use of logbook data from participating carriers to carry out a study on operational and economic impacts of the 2011 HOS regulations. ATRI has data on technology penetration among carriers, as well as video feeds from trucks and other ad hoc data for its own or government research purposes. It also collects insurance data on insurers or carriers who provide fleet-wide insurance, and this data set includes information on carriers' crash involvement and costs associated with each crash. Along with cross-sectional information, ATRI has undertaken initiatives to collect real-time data that offer the potential to address questions related to violations, scheduling, and traffic patterns.

Since 2002 ATRI, working closely with FHWA, has led the Freight Performance Measures (FPM) Program, which evaluates the effectiveness of the highway system in facilitating fast, efficient movement of goods. Performance measurements are produced for this program through the use of real, anonymous, private-sector truck data sourced through unique industry partnerships. ATRI's FPM database currently contains billions of truck data points from several hundred thousand unique vehicles spanning more than 7 years. Currently the program collects nearly 100 million data points per day, and it exceeded 1 billion points per week in late

2014. The data, which include periodic time, location, speed, and anonymous unique identification information, are collected and used by ATRI researchers to produce various indicators on truck movement, highway bottlenecks, crossing time and delay, and demand for truck routes and facilities on highways. Knowing the location of a truck or bus prior to a crash, one can estimate the number of hours driven. Also, the FPM database is a valuable source of exposure data; one can use the data to arrive at an estimate of the number of trucks on the road by time of day. However, a key limitation is that the data are collected under a strict confidentiality agreement and so are not currently available to researchers. Also, the data are collected only for a modest fraction of all trucks and buses, which may not be representative of the industries as a whole.

Data Collected by Large Truck Carriers

Most truck carriers collect information on their drivers for bookkeeping and operations management purposes. However, large carriers also often collect information on their drivers' health, wellness, crash rates, and the like. Some carriers use such data to conduct studies for purposes of improving their safety performance.

Carrier-based data can include both events and exposure, addressing the exposure problem that exists for all trucks collectively, although the data are restricted to drivers on the carrier's payroll. Also, crash data can be linked to personnel/work records, as well as to equipment manifests. Some carriers have relatively sophisticated data collection programs with respect to loss events, similar in construction to public crash files. These loss files must be used with care because they include incidents beyond police-reportable or MCMIS-reportable traffic crashes.

As part of its information-gathering process, the panel heard from Schneider National (a multinational trucking company) about the types of data it collects as part of its safety and health and wellness initiatives. This information illuminates the data elements that large truck carriers may collect on their employed drivers. Schneider collects information on crashes, which includes the time of day, the number of hours since the driver's last break, and the location of the crash and the roadway type where it occurred. The company also has electronic log data from the trucks in its fleet that track shift variability and the number of days since the truck was last at the home terminal. Data on critical events such as hard braking, roll stability control, and collision mitigation also are collected. The company prescreens its drivers for sleep apnea and treats those who test positive. As a result, the company has data not only on the safety performance of its drivers but also on their health. One outcome that potentially could be studied with these internal company data is the impact of proposed health and wellness programs on driver fatigue and safety.

OTHER DATA SOURCES

Other data sources include driver paper logs, inspection reports, and surveys.

Driver Paper Logs

Researchers exploring the relationship between driving time and the probability of crashes have collected driver logs from private fleets. In these cases, the carrier typically supplies the driver logs for a particular time period (e.g., 2 weeks), which contain information on driving patterns. This method of data collection relies on establishing collaborative agreements with private fleets. The analyses rest on valid and reliable reporting by truck drivers. A researcher obtains access to exposure data; therefore, studies based on driver logs can have a case-control formulation. As driver logs provide information on sleeper berth time and the number of rest breaks taken by a driver, researchers can investigate the impact of change in various provisions of HOS regulations provided they have data on the same set of drivers before and after such a change (Jovanis et al., 2011). The resulting analysis will be difficult to generalize given that representativeness is an issue, but the association between different driving patterns and crash risk can be estimated. Even if the study sample based on paper logs will not be random or fully representative of all truck and bus drivers, the relationship between hours of service and crash frequency may be generalizable to the CMV driver population.

Another limitation of paper driver logs is that the data are self-reported. Falsification of log books is a possibility (Moses and Savage, 1996). Monaco and Williams (2000) found that 57.8 percent of drivers in the University of Michigan Trucking Industry Program (UMTIP) 1997 data set reported driving more hours than they entered in their logbooks in the last 30 days; and a large proportion of all drivers (82.58%) said that, in general, they thought logbooks were inaccurate. Although there are no known national estimates of the prevalence of the practice of falsification of logbooks, better technology, such as electronic on-board recorders and electronic logging devices, could help address some of these quality concerns.

Inspection Reports

A state inspection system nationwide conducts more than 3 million roadside inspections of commercial motor vehicles annually to ensure that trucks and buses are operating safely. The selection of vehicles for inspection typically is not random; enforcement officers must have probable cause to inspect a vehicle. Trained inspectors in each state inspect vehicles using criteria developed by CVSA. Inspection involves an examination of the vehicle and/or driver to determine compliance with FMCSA regulations.[9] As part of the most comprehensive level I inspection, a driver's certificate from his or her medical examiner is checked, as is the driver's record of duty status and hours of service. Drivers also are checked for visible signs of fatigue. If the vehicle and/or the driver is in violation of FMCSA regulations, the vehicle and/or driver may be placed "out of service." An example of a vehicle violation is "oil and/or grease leak," while an example of a driver violation is "failing to use seat belt."

There were 3,497,937 roadside inspections in 2013 (Federal Motor Carrier Safety Administration, 2014, Table 2-5). The total numbers of vehicle and driver violations in that year were 4,118,869 and 1,047,496, respectively (as a result of some vehicles having multiple violations). Among the driver violations, 51,911 were related to driving beyond the 14-hour duty limit and 28,207 to exceeding the 11-hour driving limit. Approximately 44 percent of drivers involved in either of these types of violations were placed out of service (Federal Motor Carrier Safety Administration, 2014, Table 2-7).

Inspection reports and the MCMIS inspection file are used to identify trucking firms that are performing poorly on safety parameters. For each trucking firm, information is available on total roadside inspections, how many of its trucks were placed out of service, and the categories of the violations. This information can be used to identify sets of risk factors (at the trucking firm level) likely to characterize violators in a certain category. Moses and Savage (1996) used MCMIS roadside inspection data on 20,000 trucking firms to predict the firms' accident rates. Their empirical analysis did not control for driver characteristics. As mentioned above, data elements in the MCMIS database on the individual driver and truck/bus involved in a crash are limited; therefore, controlling for many confounding characteristics is not possible. Another limitation of inspection reports is that they include no direct measurement of driver fatigue. Nor do they provide data on vehicles that are not subject to inspection. Although not based on inspection reports, attempts have been made to identify driver- and firm-level risk factors for HOS violations using survey data. These risk factors include scheduling of irregular routes, trip lengths, compensation schemes, availability of rest areas, and type of load (Beilock, 1995; Braver et al., 1992).

---

[9] For more information on roadside inspections and inspection levels, see https://csa.fmcsa.dot.gov [March 2016].

Surveys

Useful surveys have occasionally been conducted among truck and bus drivers to gather data on their work schedules, fatigue levels, health status, access to health care services, and participation in health promotion programs. Among the surveys conducted on truck and bus drivers, two are highlighted here—one on truck drivers and the other on bus drivers—as they were most comprehensive in terms of the data elements collected.

National Survey of Long-Haul Truck Driver Health and Injury

The National Institute for Occupational Safety and Health (NIOSH) conducted a survey of long-haul truck drivers to gather baseline data on their health and safety, including the prevalence of selected health conditions and risk factors. Data were collected from 1,670 long-haul truck drivers at 32 truck stops in 20 states. Survey teams were present at each truck stop for 3 days, at varying times each day. The survey results produced estimates of the prevalence of obesity, cigarette smoking, and diabetes. The questionnaire included items on health insurance coverage and drivers' self-perception of health status. As the survey was conducted at truck stops, the truck driver population that was interviewed for the survey comprised only long-haul drivers, and excluded drivers who deliver goods locally. It is difficult to know whether the survey was strongly unrepresentative of all drivers given that there are no baseline health data on all truck drivers in the United States. The approach of recruiting truck drivers at truck stops was reasonable since it enabled the survey to reach a quasi-random population of long-haul drivers. This may be the best option available for surveying this population given its high mobility, and is preferable to interviewing drivers belonging to a specific trucking company/companies given the selection bias implicit in that approach. Mail questionnaires, phone interviews, and web-based surveys can be impractical given that CMV drivers often are away from home and have unpredictable work-rest hours.

Bus Driver Fatigue Study

The Sleep and Performance Research Center at Washington State University conducted a month-long survey from August 2010 through August 2011 of 84 commercial bus drivers (middle-aged, overweight, and predominantly male) to determine whether these drivers were working within the HOS limits set by FMCSA (for details, see Belenky et al., 2012). The bus driver population is easier to survey than the truck driver population given that most of them have regular schedules. The survey collected information on the length of various components of duty cycles. Drivers were told to keep a sleep-wake diary, and additional information on sleep-wake history was collected via actigraph. Drivers also were administered a psychomotor vigilance test (PVT) of behavioral alertness (Basner and Dinges, 2011) when they started and ended their work duty and when they went on and off on breaks during duty. Information was collected simultaneously on subjective fatigue using spontaneously perceived sleepiness and on sleepiness using the Karolinska Sleepiness Scale. The participating bus drivers represented charter (18 drivers), tour (13), regular route (25), or commuter express (24) operations; it is not known whether this percentage distribution was representative of the bus driver population.

NEEDED INFORMATION ON OPERATIONAL CHARACTERISTICS OF THE TRUCKING AND BUS INDUSTRIES

In Chapter 10, the panel argues for the need to account for all major causal factors in analysis of the causal relationship among HOS regulations, fatigue, and crash frequency. Causal factors are categorized into characteristics of the driver, the vehicle, the carrier, and the environment. This section suggests a number of important factors related to the operational characteristics of trucking for which high-quality data are not regularly collected. All these factors are important for studying driver fatigue. Collecting this information would help researchers control for confounding factors when analyzing the relationship among fatigue, hours of service, and crash frequency:

- **Exposure per hour of day:** Even though the MCMIS provides a comprehensive list of registered trucks in a particular year, the number of trucks on the nation's roads during a particular hour on a single day is unknown. This information is vital to provide a normalizing factor. Without this information, one cannot compute crash rates by time of day.
- **Trip length and driving hours:** Driving time and time on task (which encompasses driving time plus loading and unloading) are important predictors of crash risk and driver fatigue. An ATRI report on the safety impacts of HOS regulations, which consists of analysis based on TIFA data, states that 80 percent of fatal truck collisions in 2007 occurred within the first 8 hours of driving (American Transportation Research Institute, 2010, Figure 2). The importance of this percentage is difficult to assess without knowing, from exposure data, the percentage of total driving this figure represents. Figure 5-1 shows the distribution of all medium and

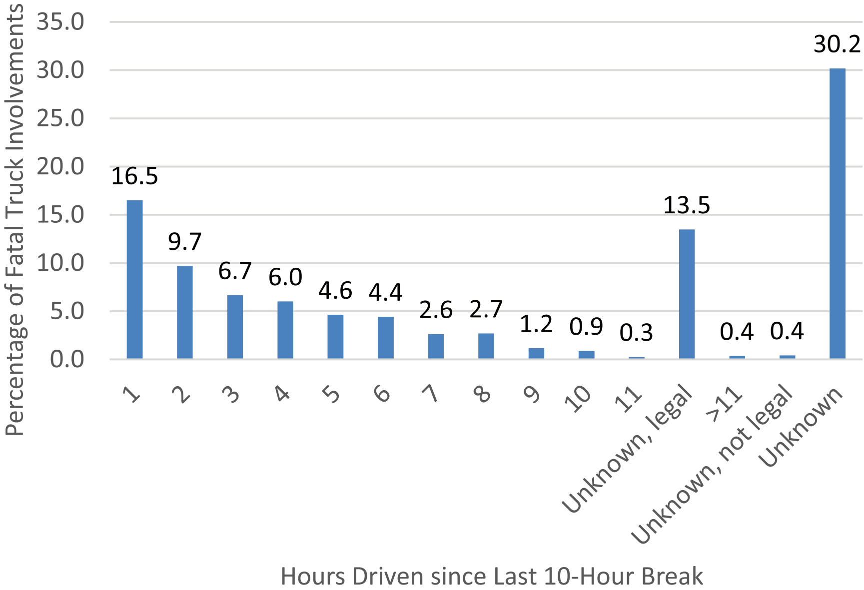

FIGURE 5-1 Truck involvement in fatal crashes by reported hours driven and year, 2004-2009.
NOTE: The data are for the period when 2003 hours-of-service (HOS) regulations were in effect.
SOURCE: Analysis by the panel using 2004-2009 Trucks Involved in Fatal Accidents (TIFA) data. Available: http://www.transtats.bts.gov/DatabaseInfo.asp?DB_ID=415&DB_Name=Trucks+Involved+Fatal+Accidents+(TIFA)&Link=0&DB_URL=Subject_ID=1&Subject_Desc=Safety&Mode_ID2=0 [March 2016].

heavy trucks involved in fatal crashes from 2004 through 2009 by driver hours behind the wheel at the time of the crash. Although the data are restricted to fatal crashes, and driver hours behind the wheel are unknown for a large percentage of fatal crashes, the figure shows a wide distribution of hours driven by truck drivers involved in fatal crashes. It is difficult to generate a distribution of hours driven by the whole population of truck drivers for comparison as that information is not available. The lack of such exposure data makes it difficult to calculate crash rates by hours driven.

- **Diversity of route and load:** Loads range from various types of freight, to liquids, to agricultural products (e.g., livestock, produce), to all sizes and types of machinery, to construction materials and equipment, to hazardous materials, to intermodal containers, to small packages, to food products, and many more. These various commodities are driven long-haul, regionally, or short-haul in rural, suburban, and urban settings. The road types and conditions vary widely across the country, and from local streets and arterials to Interstate highways.[10] Useful information about specific routes and environments might include the occurrence of various kinds of precipitation, the degree of visibility, the road surface, the roadway geometry, the degree of congestion, whether the route includes work zones, and the like.

- **Regular versus irregular schedules:** As described in Chapter 2, depending on the type of operation, drivers may have regular or irregular schedules. Irregular schedules are more likely to hamper sleep patterns and in turn affect productivity at work or safe operation of the vehicle.
- **Loading and unloading:** As discussed in Chapter 2, some drivers have loading and unloading as part of their duty, while others do not. A physically demanding task, it exposes drivers to injuries and may cut into their sleeping time. It also may increase fatigue and thus the probability of fatigued driving.
- **Operations at night (i.e., against circadian rhythm):** HOS regulations include restart and sleeper berth provisions whose purpose is to enable the driver to sleep during the nighttime (see Chapter 4). As described in greater detail in Chapter 3, human physiology makes individuals inclined to sleep during night hours, and drivers are no exception. Nighttime driving heightens the risk of fatigue.
- **Sleep quality:** Information is needed not only on hours of sleep received by the driver but also on the quality of that sleep.[11] Did the driver sleep in his or her berth or own bed at home? Sleeping in a berth may not be as comfortable and relaxing as sleeping in a bed. Darwent and colleagues (2012) investigated sleep obtained by long-haul truck drivers in Australia and found minor differences in the quality of sleep obtained in a sleeper berth versus at home.

---

[10] This report does not specifically consider off-road operations.

[11] The term “quality of sleep” is used here in the same way it is used in sleep medicine and sleep research—as the basis for a subjective complaint related to symptom reports of difficulty with sleep initiation, duration, and consolidation and daytime impairment (Buysee et al., 1989). The panel believes that the relevant scientific and medical professional sleep societies are in the best position to develop consensus definitions of the important sleep-related measures for the public. This report relies on those professional definitions and the methods for assessing them in CMV drivers relative to research, policy making, enforcement, and accident investigation.

- **Method of compensation:** Information is needed on how a driver is paid, as compensation methods can be a confounding factor. How methods and levels of compensation affect safe operations is unknown, although there is some evidence (Belzer et al., 2006) that higher levels of compensation attract and retain safer drivers.

POTENTIAL NEW DATA SOURCES

The future of motor vehicle transportation in general and commercial truck and bus transportation in particular is in flux. With the growing use of unobtrusive on-board cameras and sensors, motor vehicles soon will provide much greater assistance to the driver, and fully autonomous vehicles are now being tested. On-board technologies available today can generate extensive amounts of data related to many aspects of a vehicle, its driver, and the surrounding environment. Although these devices have the potential to identify fatigued drivers and possibly prevent them from continuing to operate their vehicle while they are impaired, liability, privacy, and security concerns will shape the future use of these technologies, as well as of the extensive real-time data they generate. Setting aside for now the privacy and confidentiality aspects of these technologies, this section briefly describes potential new data sources and the potential capabilities such data may provide to researchers in this area.

Electronic On-Board Recorders or Electronic Logging Devices

In the next few years, many carriers and owner-operators will either decide on their own or have imposed on them the obligation to carry electronic on-board recorders or electronic logging devices (ELDs) to measure, at a minimum, when and where a truck was in operation and for what duration. From a research perspective, electronic data on hours driven by a driver prior to a crash represent valuable information. Also, the expectation is that ELDs are likely to increase compliance with the HOS regulations because compared with paper logs, they are likely to be more tamper-resistant (albeit not entirely). One may speculate that they therefore could serve to reduce the extent of fatigued driving.

In 2010, FMCSA published a final rule on mandatory installation of ELDs on commercial motor vehicles manufactured after June 4, 2012. In August 2013, the Seventh Circuit Court rendered judgment that the agency could not proceed with the rule as it failed to consider driver harassment. In 2014, FMCSA proposed amendments to the rule that included requirements for the mandatory use of these devices by drivers currently required to prepare HOS records of duty status, as well as measures designed to address concerns about harassment resulting from the mandatory use of ELDs. This proposed rule is still in the comment stage. Therefore, ELDs are not currently mandatory, and the fact that industry-wide adoption of these devices will take time means that researchers using such data now will be analyzing only a selected subset of the driver population.

Telematics, On-Board Safety Systems, and Other Monitoring of Drivers

Trucks and buses increasingly are being wired for purposes of vehicle control and monitoring, as well as supervision and management of truck fleets. Telematics includes technologies for locating and tracking the vehicles (GPS) and for communicating with the driver and monitoring vehicle performance remotely. Sensors transmit a continuous stream of data from the vehicle to a central data warehouse. Depending on the kind of device installed in a truck, the vehicle's performance and condition may be captured. As the data are collected in real time, the dispatcher can warn the driver of potential problems. Vehicle data potentially available for monitoring unsafe driving include hard-braking events, sudden accelerations, and speeding. These events can be recorded for later review or communicated in real time to dispatchers. Installing such technologies is currently cost-effective for big carriers and companies that manage large fleets. The resulting data are proprietary. Data captured by telematics devices also will differ among carriers depending on their requirements.

A host of on-board safety systems are available, such as electronic stability control, roll stability control, lane departure warning, blind spot warning, forward collision warning, adaptive cruise control, and collision mitigation braking systems. These systems warn the driver of dangerous conditions, and some can be programmed to take corrective actions automatically.

On-board safety systems that use sensors and video recording systems record events outside and inside the truck. These data can be used to identify unsafe driving practices, and potentially driver fatigue. Carriers can use these data to identify at-risk drivers and develop coaching and training programs. Carriers can choose among many on-board safety systems. A recent survey conducted by UMTRI of a random sample survey of the entire fleet of trucking companies (drawn from the MCMIS) asked companies the factors that determined their choice from among a list of on-board safety systems. Companies considered the proven safety benefits of the technologies, positive feedback from drivers, driver improvement, improved safety culture, reduced cost of accidents, and insurance benefits (Belzowski et al., 2007).

Despite the lack of uniformity in the data elements trucking companies obtain from these systems, it is safe to conclude that, in combination with electronic on-board recorders, they offer the potential for a dramatic increase in the amount of driving information available, including information on hours spent driving and driving behavior.

ADVANTAGES AND LIMITATIONS OF AVAILABLE DATA SOURCES

The data sources described in this chapter have their advantages and limitations (as summarized in Table 5-2). The appropriate data source depends on the research question being pursued. No one source collects

TABLE 5-2 Advantages and Limitations of Different Data Sources

Data Source	Advantages	Limitations
Crash Databases	Provide details on driver characteristics, vehicle characteristics, road conditions, and weather conditions Generate aggregate crash statistics	Restricted to U.S. Department of Transportation (DOT)-reported crash events; lack exposure data Difficulty of identifying driver fatigue from crash reports since the data are collected by nonresearchers
Naturalistic Driving Studies	Assess driver and vehicle performance under actual road conditions Provide exposure data	Crashes are a relatively rare event; aspects of data reduction are done manually
Driving Simulators	Replicate experimental road conditions, which enables testing various scenarios Can be used to quantify performance profile of drivers who suffer from various medical conditions	Enable assessment of relative validity but not absolute validity
Electronic On-Board Recorders, Electronic Logging Devices, On-Board Safety Systems	Identify unsafe driving practices and at-risk drivers	Different set of technologies oriented toward different factors related to safety
Real-Time GPS Data	Provide exposure data	Potentially proprietary

SOURCE: Adapted from Rizzo (2011, Table 2).

comprehensive information on outcomes (crashes, noncrashes) and predictors (driver characteristics, vehicle characteristics, company characteristics, and environmental factors). Therefore, combining data from these various sources may be advantageous. The challenges entailed in doing so include the following: (1) some data sources are proprietary and would require collaboration of the data holders; (2) each data source has its own (potentially proprietary) set of definitions and taxonomies for outcomes and predictors, which would result in multiple measures of the same variable; and (3) identifiers or linking variables are very likely not available.

6

Research Methodology and Principles: Assessing Causality

One of the panel's primary tasks was to provide information to the Federal Motor Carrier Safety Administration (FMCSA) on how the most up-to-date statistical methods could assist in the agency's work. The theme of this chapter is that methods from the relatively new subdiscipline of causal inference encompass several design and analysis techniques that are helpful in separating out the impact of fatigue and other causal factors on crash risk and thereby determining the extent to which fatigue is causal.

A primary question is the degree to which fatigue is a risk factor for highway crashes. Efforts have been made to assess the percentage of crashes, or fatal crashes, for which fatigue played a key role. However, assessment of whether fatigue is a causal factor in a crash is extremely difficult and likely to suffer from substantial error for two reasons.

First, the information collected can be of low quality. Biomarkers for fatigue that can provide an objective measurement after the fact are not available. If drivers survive a crash and are asked whether they were drowsy, they may not know how drowsy they were, and even if they do know, they have an incentive to minimize the extent of their drowsiness. In most cases, the police at the scene are charged with determining whether a chargeable offense was committed; whether a traffic violation occurred; and whether specific conditions, such as driver fatigue, were or were not present. They must make this determination to the best of their abilities with limited information. It is commonly accepted and understandable that police underestimate the degree of fatigued driving and its impact on crashes.

Police assessments, augmented by more intense interviewing and other investigations, were used to determine factors contributing to crashes in such studies as the Large Truck Crash Causation Study (LTCCS) (see Chapter 5), in which the researchers attempted to determine the critical event (the event that immediately precipitated the crash) and the critical reason for that event (the immediate reason for the critical event) for each crash. To this end, they tried to provide a relatively complete description of the conditions surrounding each crash. This approach is fundamentally different from that of calculating the percentage of crashes attributable to different causes. Neither approach is entirely satisfactory: in the LTCCS approach, the concept of a "critical reason" is not well defined since many factors can combine to cause a crash, with no individual factor being solely responsible, while in the other approach, the attributed percentages can sum to more than 100 percent.

Second, in addition to low-quality information, the fact that crashes often are the result of the joint effects of a number of factors makes it difficult to determine whether fatigue contributed to a crash. Crashes can be due to factors associated with the driver (e.g., drowsiness, distractedness, anger); the vehicle (e.g., depth of tire tread, quality of brakes); the driving situation (e.g., high traffic density, presence of road obstructions, icy road surfaces, low visibility, narrow lanes); and the policies of the carrier, including its approach to compensation and to scheduling. The so-called Swiss cheese model of crash causation (Reason, 1990) posits that failures occur because of a combination of events at different layers of the phenomenon. Similarly, the so-called Haddon Matrix (Runyan, 1998) looks at factors related to human, vehicle, and environmental attributes before, during, and after a crash. A constructed matrix permits evaluation of the relative importance of different factors at different points in the crash sequence. These models acknowledge that a traffic crash has a multitude of possible causes that may not function independently, resulting in a fairly complex causal structure. Therefore, understanding the role of an individual factor, such as fatigue, in causing a crash can be a challenge.

Given that crashes can have many causes, increases and decreases in crash frequency over time can be due to changes in the frequency of any one of these causes. For instance, a harsher-than-usual winter might raise the frequency of crashes, everything else remaining constant. By ignoring such dynamics, one can be misled about whether some initiative was or was not helpful in reducing crashes.

To draw proper inferences about crash causality, then, it is important to understand and control the various causal factors in making comparisons or assessments—including those outside of one's interest, referred to as confounding factors. Therefore, to assess the degree to which fatigue increases crash risk, one must account for the dynamics of the confounding factors, including any correlation between them and the causal factors of interest. This can be accomplished through design or analysis techniques.

A common design that limits the influence of confounding factors is the randomized controlled trial. For reasons given below, however, most of the data collected in studies of motor carrier safety are observational, so methods are needed to help balance the impact of confounders on comparisons of groups with and without a causal factor of interest. By using such methods, one can better understand the role of fatigued driving and therefore help determine which policies should be implemented and warrant the allocation of resources to reduce crash risks due to fatigue.

The following sections begin by defining what is meant by causal effect. This is followed by discussion of the inferences that are possible from data on crashes and the various kinds of standardization that might be used on crash counts. Next is an examination of what can be determined through the use of randomized controlled trials and why they are not feasible for addressing many important questions. The advantages and disadvantages of data from observational studies—which are necessary for many topics in this field—are then reviewed. Included in this section is a description of techniques that can be used at the design and analysis stages to support drawing causal inferences from observational data and extrapolating such inferences to similar population groups.

DEFINITION OF CAUSAL EFFECT

The definition of a causal effect applied in this chapter is that of Rubin (see Holland, 1986). Assume that one is interested in the effect of some treatment on some outcome of interest Y, and for simplicity assume that the treatment is dichotomous (in other words, treatment or control). The potential outcome $Y(J)$ is defined as the value of the outcome Y given treatment type J. Then the *causal* effect of the treatment (as contrasted with the control) on Y_i is defined as the difference in potential outcomes $Y_i(1) - Y_i(0)$, defined as follows: a selected unit i (e.g., a person at a particular point in time) given treatment $J_i = 1$ results in $Y_i(1)$, and the same selected unit given the control $J_i = 0$ results in $Y_i(0)$, with all other factors being held constant. For example, if what would have happened to a subject under a treatment would have differed from what would have happened to the same subject at the same time under control, and if no other factors for the subject changed, the difference between the treatment and the control is said to have caused the difference. The problem when applying this definition is that for a given entity or situation, one cannot observe what happens both when $J_i = 0$ and when $J_i = 1$. One of these potential outcomes is unobserved, so one cannot estimate the unit-level causal effect. Given some assumptions about treatment constancy and intersubject independence, however, it is possible to estimate the average causal effect across a population of entities or situations. To do so, since one is comparing situations in which $J = 1$ against those in which $J = 0$, one must use techniques that make it possible to assert that the units of analysis are as similar as possible with respect to the remaining causal factors.

Understanding causality is an important goal for policy analysis. If one understands what factors are causal and how they affect the outcome of interest, one can then determine how the changes to causal factors even for a somewhat different situation from the one at hand will affect the probability of various values for the outcome of interest. If one simply determines that a factor is associated with an outcome, however, it may be that the specific circumstances produced an apparent relationship that was actually a by-product of confounding factors related to treatment and outcomes.

DRAWING INFERENCES AND STANDARDIZING CRASH COUNTS

As one example of confounding and the challenges entailed in drawing causal inferences, it is common for those concerned with highway safety to plot crash counts by year to assess whether road safety is improving for some region. This type of analysis can be misleading. For example, Figure 6-1 shows a large decline in total fatalities in truck crashes between 2008 and 2009. It is generally accepted that this decline was due to the substantial reduction in vehicle-miles traveled that resulted from the recession that started during that year. However, it is also possible that the decline was due in part to new safety technology, improved brakes, improved structural integrity of the vehicles, or increased safety belt use. Thus, looking at a time series of raw crash counts alone cannot yield reliable inferences.

As a first step in enabling better interpretation of the data, one could standardize the crash counts to account for the change in vehicle-miles traveled, referred to as exposure data. Thus an obvious initial idea is to use vehicle-miles traveled as a denominator to compute crashes or fatal crashes per miles traveled. In some sense, exposure data are a type of confounding factor, because a truck or bus that is being driven less is less likely to be involved in a crash. The lack of exposure data with which to create crash rates from the number of crashes is a problem discussed below. Another problem with normalizing crashes by dividing by vehicle-miles traveled is that the relationship between the number of crashes and

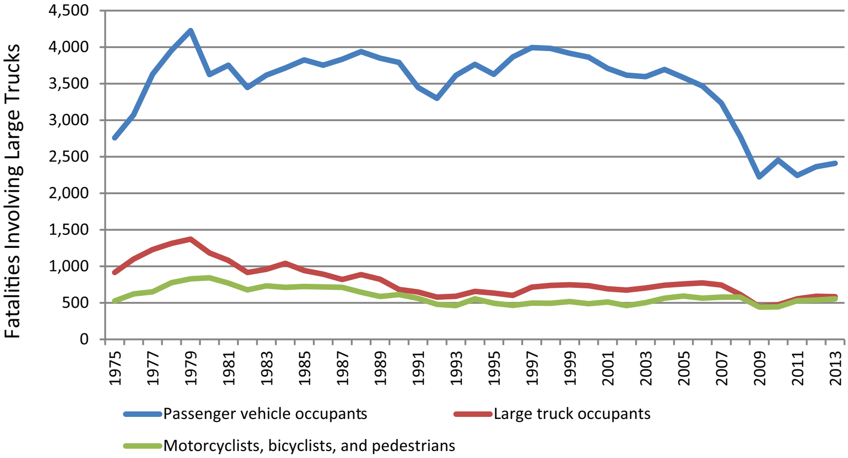

FIGURE 6-1 Deaths in crashes involving large trucks, 1975-2013.
SOURCE: Insurance Institute for Highway Safety. Available: http://www.iihs.org/iihs/topics/t/large-trucks/fatalityfacts/large-trucks [March 2016] based on the U.S. Department of Transportation's Fatality Analysis Reporting System.

the amount of exposure might be nonlinear, as pointed out by Hauer (1995). This nonlinearity is likely due to traffic density as an additional causal factor.

The idea of standardization can be extended. What if other factors could confound the comparison of time periods? For example, suppose that in comparing two time periods, one finds that more miles were traveled in 1 year under wet conditions than in the other year? To address this potential confounder, the data could be stratified into days with and without precipitation prior to standardizing by vehicle-miles traveled. Increasingly detailed stratifications can be considered if the data exist for various factors. Yet there are limits to which this can be done. At some point, one would have such an extensive stratification that there would likely be few or no crashes (and possibly even no vehicle-miles traveled) for many of the cells. To address that issue, modeling assumptions could be used in conjunction with various modeling approaches. For instance, one could assume that log $[\Pr(Crash)/(1 - \Pr(Crash))]$ is a linear function of the stratifying factors, but this approach would rely on these assumptions being approximately valid.

An understanding of which factors are and are not causal and the extent to which they affect the outcome of interest is important in deciding on an appropriate standardization. Efforts at further standardization by other potential casual factors or potential confounders are likewise constrained by the fact that police reports often include only limited information on the driver, the vehicle, and the environment.

At present, the main source of data for vehicle-miles traveled is the Federal Highway Administration (FHWA). However, these data are too aggregate and lacking in specifics to be used as denominators in producing crash rates for various kinds of drivers, trucks, and situations. Without exposure data, one might be able to separate collisions into those in which a factor was or was not present (although doing so is difficult, see Chapter 5). However, since one would not know how much crash-free driving had occurred when that factor was and was not present, one could not know whether the number of crashes when a factor was present was large or small.

ROLE OF RANDOMIZED CONTROLLED TRIALS

Much of what is known about what makes a person drowsy, how being drowsy limits a one's performance, and what can be done to mitigate the effects of inadequate sleep derives from laboratory studies, which commonly entail randomized controlled trials. For instance, studies have been carried out with volunteers to see how different degrees of sleep restriction affect response time. For such an experiment, it is important for the various groups of participants to differ only with respect to the treatment of interest—for example, degree of sleep restriction—and for them not to differ systematically on any confounding factors. In randomized experiments, one minimizes the effects of confounders by randomly selecting units into treatment and control groups. As the sample size increases, the randomization tends to balance all confounders across the different groups. (That is, randomization causes confounders to be uncorrelated with selection into treatment and control groups.) Traditional randomized controlled trials also are usually designed to have relatively homogeneous participants so that the treatment effect can more easily be measured. This homogeneity is achieved by having restrictive entry criteria. Further, the treatment is usually constrained as well. While this homogeneity of participants and intervention improves assessment of the efficacy of the treatment effect, it often limits the generalizability of the results.

In addition to restrictive entry criteria, stratification or matching is used to provide greater control over potential confounding characteristics. If such techniques are not used, the result can be an imbalance between the treatment and control groups on such characteristics, even with randomization into groups. For example, one could have more elderly people in the treatment group than in the control group even with randomization. As the number of potentially causal factors increases, the opportunities for such imbalance also increase.

As discussed below, for a number of topics involving field implementation, randomized controlled trials are not feasible. One type of study, however—the randomized encouragement design—provides some of the benefits of such trials but may be more feasible. In such studies, "participants may be randomly assigned to an opportunity or an encouragement to receive a specific treatment, but allowed to choose whether to receive the treatment" (West et al., 2008, p. 1360). An example would be randomly selecting drivers to receive encouragement to be tested for sleep apnea and examining the effects on drivers' health (following Holland, 1988). This type of design can be useful when the treatment of interest cannot be randomly assigned, but some other "encouragement" to receive the treatment (such as a mailer or monetary incentive) can be randomly provided to groups of participants.

Before continuing, it is important to reiterate that current understanding of the influence of various factors on highway safety and on fatigue comes from a variety of sources, including laboratory tests, naturalistic driving studies, and crash data (see Chapter 5). These various sources have advantages and disadvantages for addressing different aspects of the causal chain from various sources of sleep inadequacy, including violation of hours-of-service (HOS) regulations, to sleep deficiency, to lessened performance, to increased crash risk. One can think of these various sources of information as being plotted on a two-dimensional graph of fidelity versus control. Typically, as one gains fidelity—that is, correspondence with what happens in the field—one loses control over the various confounding factors. That is why it can be helpful to begin studies in the laboratory, but as one gains knowledge, some field implementation is often desirable. These latter studies will often benefit from methods described in the next section for addressing the potential impacts of confounding factors.

OBSERVATIONAL STUDIES

Observational studies are basically surveys of what happened in the field (e.g., on the road). If data were gathered from individuals who did and did not receive some intervention or treatment or did and did not engage in some behavior, one could compare any outcome of interest between those groups. However, any such comparison would suffer from a potential lack of comparability of the treatment and control groups on confounding factors. That is why techniques are needed to help achieve such balance after the fact. However, observational studies do have the advantage of collecting data that are directly representative of what happens in the field.

Further, such studies are generally feasible, which often is not the case for randomized controlled trials. For example, it is not possible to randomize drivers to follow or not follow the HOS regulations. Such an experiment would obviously be unethical as well as illegal. Similarly, drivers diagnosed with obstructive sleep apnea could not be randomly divided into two groups, one treated with positive airway pressure (PAP) devices and the other not, to assess their crash risk on the highway. For most issues related to study of the role of fatigue in crashes, such random selection into treatment and control groups is not feasible.

With a few exceptions, the data currently collected that are relevant to understanding the linkage between fatigue and crash frequency are observational (nonexperimental). Therefore, methods are needed for balancing the other causal factors between two groups that differ regarding some behavior or characteristic of interest so those other factors will not confound the estimates of differences in that factor of interest. For example, not properly controlling for alcohol use may lead to an overestimation of the effects associated with fatigue for nighttime driving. Thus without careful design and analysis, what one is estimating is not the effect of a certain factor on crash frequency but the combination of the effect of that factor and the difference between the treatment and control groups on some confounding factor(s).

This point is illustrated by a study undertaken recently by FMCSA to determine whether the method of compensation of truck drivers is related to crash frequency. Here the type of compensation is the treatment, and crash frequency is the outcome of interest. A complication is that carriers who chose a specific method for compensation might have other characteristics over- or underrepresented, such as their method for scheduling drivers or the type of roads on which they travel. It is difficult to separate the effect of the compensation approach from these other differences among carriers.

Regression Adjustment

Instead of balancing these other causal factors by matching or stratifying, one might hope to represent their effect on the outcome of interest directly using a regression model. Here the dependent variable would be the outcome of interest, the treatment indicator would be the primary explanatory variable of interest, and the remaining causal factors would be additional explanatory variables. The problem with this technique is that the assumption that each of the explanatory variables (or a transformation of a variable) has a specific functional relationship with the outcome is a relatively strong assumption that is unlikely to be true. The farther apart are the values for the confounding factors for the treatment and control groups, the more one will have to rely on this assumption. (There are also nonparametric forms of regression in which the dependence on linearity is reduced, but some more general assumptions still are made about how the outcome of interest and the causal factors interact, for example, see Hill [2011].)

Design Methods for Observational Data

This section describes three techniques used in conjunction with the collection of observational data in an attempt to derive some of the benefits of a randomized controlled trial by limiting the influence of confounding factors. Note that this is an illustrative, not a comprehensive list, and the terminology involved is not altogether standardized.

Cohort Study

A cohort of cases is selected and their causal factors measured as part of an observational study database. Then either the cases are followed prospectively to ascertain their outcome status, or that assessment is performed on historical records as part of a retrospective study.

Case-Control Study

To assess which factors do and do not increase the risk of crashes, one can identify drivers in an observational database who have recently been involved in crashes, and at the same time collect information on their characteristics for the causal factor(s) of interest and for the confounding causal factors. Then, one identifies controls that match a given case for the confounding factors from among drivers in the database who have not been involved in recent crashes. One next determines whether the causal factor(s) of interest were or were not present more often in the cases than in the controls. An example might be to see whether fewer of those drivers recently involved in a crash relative to controls worked for a safety-conscious carrier, controlling for the driver's body mass index (BMI), experience, and other factors. If one did not match the two groups of drivers on the confounding factors, this approach could produce poor inference, since the two groups likely would differ in other respects, and some of those differences might be causal.

Case-Crossover Study

A case-crossover design is used to answer the question: "Was the event of interest triggered by some other occurrence that immediately preceded it?" (Maclure and Mittleman, 2000; Mittleman et al., 1995). Here, each case serves as its own control. The design is analogous to a crossover experiment viewed retrospectively. An example might be a truck driver who had been involved in a crash. One might examine whether the truck driver had texted in the previous hour and then see whether the same driver had texted a week or a month prior to the crash, and again for several previous time periods. In that way, one would obtain a measure of exposure to that behavior close to the time of the crash and exposure more generally. (Of course, assessing whether a driver has texted is not always straightforward.)

Analysis Methods for Observational Data

This section describes some analytic methods that can be used to select subjects for analysis or to weight to achieve balance between a treatment and a control group on confounding factors.

Propensity Score Methods

One of the most common tools for estimating causal effects in nonexperimental studies is propensity score methods. These methods replicate a randomized experiment to the extent possible by forming treatment and comparison groups that are similar with respect to the observed confounders. Thus, for example, propensity scores would allow one to compare PAP device users and nonusers who appear to be similar on their prestudy health behaviors, conditions, and driving routines. The propensity score summarizes the values for the confounders into the propensity score, defined as the probability of receiving treatment as a function of the covariates. The groups are then "equated" (or "balanced") through the use of propensity score weighting, subclassification, or matching. (For details on these approaches, see Rosenbaum and Rubin [1983]; Rubin [1997]; and Stuart [2010]. For an application of this method to highway safety, see Wood et al. [2015].)

Propensity score methods utilize a model as does regression adjustment, but not in the same way. Propensity score methods have two features that provide an advantage relative to regression adjustment: (1) they involve examining whether there is a lack of overlap in the covariate distribution between the treatment and control groups, and whether there are certain values of the covariates at which any inferences about treatment effects would rely on extrapolation; and (2) they separate the design from the analysis and allow for a "blinded approach" in the sense that one can work hard to fit the propensity score model and conduct the matching, weighting, or subclassification (and assess how well they worked in terms of balancing the covariates) without looking at the outcome.

Both propensity score methods and regression adjustment rely on the assumption that there are no unmeasured confounding factors. Techniques described below, such as instrumental variables and regression discontinuity, are ways of attempting to deal with potential unmeasured confounding. The assumption of no unmeasured confounders cannot be tested, but one can use sensitivity analyses to assess how sensitive the results are to violations of that assumption (for details, see Hsu and Small [2013]; Liu et al. [2013]; and Rosenbaum [2005]).

Marginal Structural Models

Propensity score methods are easiest to use when there is a relatively simple and straightforward time ordering: (1) a point-in-time treatment with covariates measured before treatment, (2) a treatment administered at a single point in time, and (3) outcomes measured after treatment. For more complex settings with time-varying covariates and treatments, a generalization of propensity score weighting—marginal structural models—can be used (for details, see Cole and Hernan [2008] and Robins et al. [2000]). These approaches are useful if, for example, one has data on drivers' PAP use over time, as well as on measures of their sleep or health status over time, and one wants to adjust for the confounding of health behaviors over time.

The basic idea of the marginal structural model is to weight each observation to create a pseudopopulation in which the exposure is independent of the measured confounders. In such a pseudopopulation, one can regress the outcome on the exposure using a conventional regression model that does not include the measured confounders as covariates. The pseudopopulation is created by weighting an observation at time t by the inverse of the probability of the observation's being exposed at time t, that is, by weighting by the inverse of the propensity score at time t.

As noted, marginal structural modeling can be thought of as a generalization of propensity score weighting to multiple time points. To describe the method informally, at each time point, the group receiving the intervention (e.g., those receiving PAP treatment at that time point) is weighted to look similar to the comparison group (those not receiving PAP treatment at that time point) on the basis of the confounders measured up to that time point. (These confounders can include factors, such as sleep quality, that may have been affected by a given individual's prior PAP use). As in propensity scoring, the weights are constructed as the estimated inverse of probability of receiving the treatment at that point in time. So those individuals who have a large chance of receiving the treatment are given a smaller weight, and similarly for the comparison group, which results in the groups being much more comparable. The causal effects are then estimated by running a weighted model of the outcome of interest (e.g., crash rate) as a function of the exposure of interest (e.g., indicator of PAP use). (The measured confounders are not included in that model of the outcome; this is known as the "structural" model).

Use of Multiple Control Groups

Using multiple control groups is a way of checking for potential biases in an observational study (Rosenbaum, 1987; Stuart and Rubin, 2008). An observational study will be biased if the control group differs from the treatment group in ways other than not receiving the treatment. In some settings, one can choose two or more control groups that may have different potential biases (i.e., may differ from the treatment group in different ways). For example, if one wanted to study the annual change in crash rates due to truck drivers' having increased their BMI by more than 5 points in the previous year to a total of more than 30, such truck drivers might be compared with drivers who had BMIs that had not changed by more than 5 points and still had BMIs under 30, and the same for bus drivers. If the results of these comparisons were similar (or followed an expected ordering), the study findings would be strengthened. Thus, for example, the findings would be stronger if one of the two control groups differed in that one had a higher expected level of unmeasured confounders than the treatment group had, while the other control group had a lower expected level, and the results were consistent with that understanding. If, however, one believed that there were no unmeasured confounders, but the control groups differed significantly from each other (so that the comparisons of the treatment and control groups differed significantly), that belief would have to be wrong, since the difference in control groups could not be due to the treatment. (This is referred to as bracketing and is described in Rosenbaum [2002, Ch. 8].)

Instrumental Variables

Another common technique for use with observational data is instrumental variables. This approach relies on finding some "instrument" that is related to the treatment of interest (e.g., the use of some fatigue alerting technology) but does not directly affect the outcome of interest (e.g., crash rates). In the fatigue alerting example, such an instrumental variable could be the indicator of a health insurance plan that provides free fatigue alerting devices to drivers. Drivers in that plan could be compared with those not in the plan, under the assumption that the plan might increase the likelihood of drivers using such a device but would not directly affect their crash risk, except through whether they used the device. The advantage here is that there would be a good chance that the drivers who did and did not receive the free devices would be relatively comparable (possibly depending on additional entry criteria for the program).

The introduction of such instrumental variables can be a useful design, but it can be difficult to identify an appropriate instrumental variable that is related strongly enough to the treatment of interest and does not have a direct effect on the outcome(s) of interest. One potentially useful approach to addressing this issue is use of an encouragement design (similar to that discussed above), in which encouragement to receive the treatment of interest is randomized. Using PAP devices as an example, a randomly selected group of drivers would be given some kind of encouragement to use the devices. This randomized encouragement could then be used as an instrumental variable for receiving and using the device, making it possible to examine, for example, the effects of PAP use on crash rates. (For more examples of and details on instrumental variables, see Angrist et al. [1996]; Baiocchi et al. [2010]; Hernán and Robins [2006]; and Newhouse and McClellan [1998].)

Regression Discontinuity

Regression discontinuity can be a useful design when an intervention is administered only for those exceeding some threshold quantity. For example, everyone with a hypopnea score above some threshold would receive a PAP device, and those below the threshold would not. The analysis then would compare individuals just above and just below the threshold, with the idea that they are likely quite similar to one another except that some had access to the treatment of interest while others did not. Bloom (2012) provides a good overview of these designs.

Interrupted Time Series

Interrupted time series is a useful approach for estimating the effects of a discrete change in policy (or law) at a given time (see, e.g., Biglan et al., 2000). The analysis compares the outcomes observed after the change with what would have been expected had the change not taken place, using data from the period before the change to predict that counterfactual.

One useful aspect of this approach is that it can be carried out with data on just a single unit (e.g., one state that changed its law), with repeated observations before and after the change. However, the design is stronger when there are also comparison units that did not implement the change (such as a state with the same policy that did not change it), which can help provide data on the temporal trends in the absence of the change. This could be useful, for example, for examining the effect of a change in a company health program if data also were available from a company that did not make the change at that time. These designs, with comparison subjects, are known as "comparative interrupted time series" designs.

A special case of comparative interrupted times series is difference-in-difference estimation, which is basically a comparative interrupted time series design with only two points, before and after the change. This approach compares the differences before and after the change between two groups, one that did and one that did not experience the change. This approach enables controlling for secular changes that would have taken place in the absence of the change of interest, as well as differences between the groups that do not change over time. (A good reference for these designs is Meyer [1995].)

Sensitivity Analysis

For propensity score approaches, instrumental variable analyses, and many of the other techniques described here, it is useful to determine the robustness of one's inference through the use of sensitivity analysis. As noted above, one of the key assumptions of propensity score matching is that bias from unobservable covariates can be ignored. If one could model the effect of unobserved covariates, one could test this assumption by calculating the difference between estimated treatment effects—after controlling for observed covariates and the effect of unobserved covariates. If the estimated treatment effect were essentially erased by unobservable covariates, one could conclude that the treatment effect was due to the bias from unobservable covariates and was not a true treatment effect. However, testing the assumption is impossible because researchers do not have data on unobservable covariates. Therefore, a researcher would need to obtain a proxy for the bias from unobserved covariates, which would require a detailed understanding of the phenomenon being researched. As a result, sensitivity analysis procedures involve examining how much unmeasured confounding would need to be present to alter the qualitative conclusions reached and then trying to determine whether that degree of confounding is plausible. (For details, see Hsu and Small [2013]; Liu et al. [2013]; and Rosenbaum [2005].)

Generalizing Findings from Observational Studies to a Different Population

Often it is necessary to draw inferences for a population for which directly relevant research has not been carried out. A key example in the present context is drawing inferences about commercial motor vehicle drivers when the relevant research is for passenger car drivers. When is it safe to make such an extrapolation?

In this case, one needs first to assess internal validity for the population on which the relevant study was done, and then assess the generalizability of the findings to another population of interest. The internal validity question involves the strength of the causally relevant inference that can be drawn about a given research question for the population and treatment studied (which may differ from the population and treatment of interest). The answer will naturally depend on the study design and analysis plans. Different study designs have different implications regarding what can be concluded. The second issue is the generalizability of the findings. The hope is that the findings can be translated to the administration of the same or a closely related treatment for a similar population.

Criteria for determining the degree to which a study enables causal inference have been considered for many decades. In the area of medical and epidemiologic studies, one well-recognized set of criteria was advanced by Hill (1965). These criteria have evolved over time, and a summary of their modern interpretation is as follows:

- Strength of association between the treatment and the outcome: The association must be strong enough to support causal inference.
- Temporal relationship: The treatment must precede the outcomes.
- Consistency: The association between treatment and outcomes must be consistent over multiple observations among different populations in different environments.
- Theoretical plausibility: There must be a scientific argument for the posited impact of the treatment on the outcome.
- Coherence: The pattern of associations must be in harmony with existing knowledge of how the treatment should behave if it has an effect.
- Specificity: A theory exists for how the treatment affects the outcome of interest that predicts that the treatment will be associated with that outcome in certain populations but not associated (or less associated) with other outcomes and in other populations, and the observed associations are consistent with this theory. Furthermore, alternative theories do not make this same set of predictions (Cook, 1991; Rosenbaum, 2002).
- Dose-response relationship: Greater exposure to the risk factor is associated with an increase in the outcome (or a decrease if the treatment has a negative effect on the outcome).
- Experimental evidence: Any related research will make the causal inference more plausible.
- Analogy: Sometimes the findings can be extrapolated from another analogous question.

The panel suggests an additional criterion—the elimination of alternative explanations for observed associations—to be key in helping to establish a causal relationship.

While criteria for establishing causal relationships have evolved over time, the principles articulated by Hill are still valid. The panel wishes to emphasize the criteria of consistency, theoretical plausibility, coherence, and experimental evidence, which support the point that causal inference often is not the result of a single study but of a process in which evidence accumulates from multiple sources, and support for alternative explanations is eliminated. As described in this chapter, the past 30 years also have seen many advances regarding methods for estimating the effects of "causes" or interventions in nonexperimental settings.

There is value, then, in using a variety of approaches to better understand the arguments that can be made as to whether a treatment or an intervention has an effect. Doing so makes it possible to gain causally relevant knowledge from the collection of relevant studies so as to obtain the best possible understanding of the underlying phenomenon.

A good example of how causality can be established primarily through observational studies is the relationship between cigarette smoking and lung cancer. In the 1950s, Doll and Hill (1950) and others carried out a number of observational studies on the association between cigarette smoking and lung cancer. These studies had the usual limitations and potential for confounding factors common to such studies. Yet strong associations were found across multiple populations and settings, and this association also was shown to be monotonically related to the amount of smoking (see Hill's criterion on the dose-response relationship above). Some, however, including R.A. Fisher, proposed an alternative explanation: that there existed a factor that increased both the likelihood a person would use tobacco and the risk of contracting lung cancer, such as a genetic variant that made a person more likely to smoke and more likely to contract lung cancer through independent mechanisms. This alternative hypothesis was placed in doubt by a sensitivity analysis showing that if such a factor existed, it would need to have an association with smoking at least as great as the observed association between smoking and lung cancer, and the proposed factors, such as genetic variants, were unlikely to have such a strong association with smoking. Other alternative hypotheses were systematically rejected (see Gail, 1996). Even though a randomized controlled study of tobacco use was clearly infeasible, it became clear through the variety of available studies that supported the hypothesis and failed to support the rival hypotheses that cigarettes were a causal factor for lung cancer.

The spectrum of observational study types includes retrospective cohort and case-control studies, prospective studies, and various types of designs based on observational data, described by Shadish and colleagues (2002). These techniques, and additional ideas described here, have been applied in a number of policy areas and can be used to reduce the opportunity for confounding factors to influence outcomes when a study does not have a randomized controlled design.

Once treatment efficacy has been addressed through a causal understanding of the phenomenon, one is left with the question of the generalizability of the findings from the available studies to other populations and interventions. What one would like is to have a sufficiently clear understanding of the science underlying a finding of treatment efficacy that one can transfer the finding to the administration of the same or a closely related treatment for different populations. For an excellent discussion of this issue, see Pearl and Bareinboim (2011).

PART III

CURRENT RESEARCH FINDINGS

7

Fatigue, Hours of Service, and Highway Safety

It is well known from laboratory studies that fatigue can cause shortfalls in performance, including slower response times, attention failures, and poor decision making (see Chapter 3). It also is well known from empirical data collection that fatigue can result in an increased risk of crashes, which is due to these decreases in performance. Further, it is reasonable to believe that restrictions on hours of service lead to a reduction in the percentage of fatigued drivers. However, this linkage is complicated by other dynamics that argue against such a simple causal statement. Given the current hours-of-service (HOS) regulations, there would be primary interest in assessing the difference in fatigue between driving for 10 hours and driving for 11 hours in a day, since that is the current policy-relevant question relative to setting HOS regulations. This causal question is framed somewhat narrowly, and the panel believes it is important to view this issue more broadly for the following reason. The nation is experiencing increased use of technology in vehicles that could be effective for fatigue detection (see Chapters 9 and 11), and other improvements are being made in the design of trucks and buses, in the driving environment, possibly in commercial motor vehicle (CMV) drivers' personal habits, and in the scheduling policies of carriers. Therefore, the question of greatest importance is what factors mediate the above causal path from fatigue to performance shortfalls to crashes, and therefore how accidents can be prevented in a variety of future circumstances, not simply those being faced today.

For present purposes, then, the panel is interested in whether this causal chain is in operation for truck and bus drivers and in the characteristics of its operation. The laboratory and empirical evidence for passenger cars is clear, but given the few rigorous studies for truck and bus drivers and their different driving circumstances, it remains possible that truck and bus drivers, as professionals, may be different. For instance, they may be better able to judge when their performance has been compromised and take some action in response to reduce the risk. Such action might include pulling over and resting; consuming caffeine; or in the longer term, changing one's sleeping habits or a carrier's changing the scheduling of drivers. In addition, when trying to address causal factors such as fatigue in a system as complex as CMV driving, one must be aware of the possibility of unintended consequences. For example, the existence of rumble strips may give drivers a false sense of security and so may encourage them not to pull over. Also, having restart provisions that require no driving between specified ranges of hours may increase the density of driving just outside the boundaries. For this reason, it may be useful to examine the literature on sociotechnical systems, which can be helpful in identifying such unintended consequences (Carayon, 2006; Hanowski, 2013).

After providing an introduction to crash risk due to fatigue, this chapter summarizes the techniques that have been used and the data sets to which they have been applied in some of the leading research on how increases in hours of service and increases in fatigue are linked to increases in crash risk for CMV drivers. There is also a relevant literature on the factors that underlie fatigue in CMV drivers. The review in this chapter is not meant to be comprehensive but is based on a selection of key reports that highlight the methods and data used in this research and the kinds of results such methods provide. Given the large number of confounding factors involved (see Table 10-1 in Chapter 10), which include health and other characteristics of the driver, the vehicle used, the driving environment, and policies of the carrier, along with the difficulty of collecting data on many of these factors, it should not be surprising that much remains to be learned about the relationship among hours of service, fatigue, and crash risk. One of the key research issues is the need to separate out the impacts of hours worked, time of day, and number of hours slept to determine the extent to which each affects fatigue and therefore crash risk. This kind of analysis will require more sophisticated statistical models than have routinely been applied in this area. Some of these techniques are discussed in Chapter 6, and Chapter 10 provides a conceptual approach for moving forward.

This chapter concludes by identifying questions that need additional research. The hope is to both help policy makers understand the complexity of the issue and to help guide future researchers in deciding where to focus their efforts to reduce the remaining uncertainties.

It should be noted that all but three of the studies included in the literature review in this chapter (Connor et al. [2002], Guo et al. [2010], and Tefft [2010, 2014]) involved data on commercial motor vehicles rather than passenger cars. The panel understands that there exists an extensive literature on passenger car crashes and the relationship between crash risk and fatigue for automobiles that these three sources only touch on. The reason for the deemphasis on this related literature in the present review is that the causal model for truck and bus drivers is likely considerably different from that for passenger car drivers with respect to the strength of the individual causal factors that need to be accounted for, but perhaps also with regard to which causal factors are involved. Relative to driving a passenger vehicle, CMV driving often involves longer periods of continuous driving, greater fractions of a day and of a week spent driving, the resulting lifestyle, the economic pressures to continue driving when fatigued, the physical demands of loading and unloading, and the differences in driving a truck or bus compared with a passenger car, not to mention the lack of an analogue to HOS regulations. All of these factors contribute to the panel's view that the emphasis here should be on research on the relationship among fatigue, hours of service, and crash risk for CMV drivers.

INTRODUCTION TO ASSESSMENT OF THE ROLE OF FATIGUE IN INCREASING THE RISK OF TRUCK AND BUS CRASHES

In the United States, 3,921 people were killed and 104,000 were injured in crashes involving large trucks in 2012. The analogous statistics for bus crashes for 2011 were 283 and 24,000, respectively. Crash databases compiled from police reports, such as the Fatality Analysis Reporting System (FARS), are sometimes used to provide estimates of the number of crashes involving trucks or buses that were associated with a fatigued driver. In particular, FARS has been used to estimate that 1.5 percent of crashes involving large trucks that resulted in a fatality in 2013 were due to the truck driver being asleep or fatigued.

Many view this estimate, and similar estimates for nonfatal crashes, as biased low because driver fatigue is difficult to detect during police accident investigations (the basis for FARS reporting) (see Chapter 5). Additionally, police investigators, not usually trained in how to recognize fatigue post hoc, are somewhat reluctant to identify it as such on crash reports because they subsequently will be expected to explain in court why they labeled a crash as related to driver fatigue. If a vehicle is not instrumented for the purpose, it is extremely difficult to determine whether fatigue contributed to a crash. And as argued in Chapter 5, even if a vehicle is instrumented, there remain situations in which it is unclear whether fatigue was a primary contributing factor.

An often-quoted, relatively high estimate of the percentage of fatal truck crashes associated with a fatigued driver resulted from a 1990 report by the National Transportation Safety Board (NTSB) (National Transportation Safety Board, 1990). As discussed in greater detail below, the NTSB conducted an in-depth examination of 182 crash reports on fatal-to-the-driver large-truck crashes that occurred in eight states between October 1, 1987, and September 30, 1988. The analysis considered information available on the number of hours recently driven, the type of accident (such as colliding with a vehicle ahead of the truck or gradually veering out of one's lane), and time of day. Driver fatigue was identified as a principal cause in 31 percent (56 of 182) of those fatal-to-the-driver crashes (National Transportation Safety Board, 1990). Indeed, of all the factors investigated, driver fatigue turned out to be the most frequent cause. The NTSB estimated at the time that the sample of crash reports it examined in depth represented about 25 percent of such fatal-to-the-driver reports nationwide.

Since fatal-to-the-driver crashes are a subset of fatal crashes, which in turn are a subset of crashes involving trucks (or buses), this percentage should not be applied to larger crash populations. (For example, Knipling and Wang [1994] estimated that 1-4 percent of truck crashes were related to driver fatigue.)

A more realistic estimate of the percentage of serious truck crashes linked to driver fatigue comes from the Large Truck Crash Causation Study (LTCCS), conducted by the Federal Motor Carrier Safety Administration (FMCSA) and the National Highway Traffic Safety Administration (NHTSA) between 2001 and 2003. This study entailed conducting in-depth investigations of 963 crashes involving a truck that resulted in a fatality or an injury to determine the critical reasons for these crashes, each of which was assigned to one or more reasons (see below and Chapter 5 for more details on this study). Truck driver fatigue was found to be *associated* with 13 percent of the crashes (Starnes, 2006). This means that one of the drivers involved was found to be fatigued, but it was not established whether that fatigue was an important contributor to the crash.

There also has been some international work on this issue. In England, for example, Horne and Reyner (1995) found that 16 percent of all vehicle crashes were related to sleep insufficiency. More recently, Garbarino and colleagues (2001) determined that 21.9 percent of highway crashes in Italy between 1993 and 1997 were related to sleepiness. While these studies are not specific to truck crashes, they provide some indication of the importance of driver fatigue as a cause of crashes.[1]

---

[1] Another study (Stevenson et al., 2014) was brought to the panel's attention when this report was nearly finalized. It is a careful case-control study of the causes of nonfatal, nonsevere crashes involving heavy vehicles in Australia.

The wide range of estimates of the degree to which sleepiness or fatigue can be shown to be associated with truck collisions is due in part to the many differences in what is being estimated. These differences include the degree of severity of the crash, whether cars or trucks and buses were involved, the time and location of data collection, and the definition of operator fatigue. In the end, despite the various attempts that have been made to estimate the incidence of truck driver fatigue contributing to crashes on the nation's roadways, the panel simply did not find convincing enough evidence that at this time would support a reliable estimate.

A literature review conducted by Belenky and colleagues (2013) to assess the prevalence of fatigue in bus drivers and its association with crash risk found that none of the studies included in the review investigated the impact of nonpathologic fatigue on the driving ability of these drivers. Putcha and colleagues (2002) analyzed FARS data looking for fatal crashes in which bus drivers were involved during the period 1995-1999. The FARS database includes the contributing factor "drowsy, asleep, or fatigued," so the authors were able to determine the fatal crashes in which this factor was indicated as occurring. They found this to be the case for only 5 of a total of 1,483 fatal crashes involving buses over the 5-year period. The panel believes that the resulting estimate of 0.3 percent is almost certainly an underestimate, for the reasons discussed earlier. Clearly, this is an area in need of further research.

RESEARCH ON FATIGUE, HOURS OF SERVICE, AND RISK OF COMMERCIAL MOTOR VEHICLE CRASHES

Following are summaries and critiques of some of the key research examining the relationship among CMV driver fatigue, HOS regulations, and crash risk. For further information on most of these studies, see Knipling (2015) at http://sites.nationalacademies.org/DBASSE/CNSTAT/CMV_Driver_Fatigue_Long-Term_Health_and_Highway_Safety/index.htm.

Crash Involvement of Large Trucks by Configuration: A Case-Control Study

(Stein and Jones, 1988)

This was a prospective case-control study of the causes of large-truck crashes. According to the authors, "For a two-year period, large truck crashes on the interstate system in Washington State were investigated using a case-control method. For each large truck involved in a crash, three trucks were randomly selected for inspection from the traffic stream at the same time and place as the crash but one week later. The effects of driver and truck characteristics on crash risk was assessed by comparing their relative frequency among the crash-involved and the comparison sample trucks." The data set represented 676 crashes involving 734 trucks that occurred between 1984 and 1986. The characteristics assessed for crash and control trucks were truck configuration, age of driver, weight of load, hours of driving, truck body type, and fleet size. The trucks also were inspected to check on the condition of the brakes, steering, and tires. (Continuous variables were classified into three groups of equal size to define matching cases.) To determine whether a variable was distributed differently in the crash versus the control population, the percentage of trucks with that characteristic in the crash population was divided by the percentage of trucks with that characteristic in the control population. In addition, to deal with simultaneous effects of the various characteristics, a logistic regression model was used to estimate the adjusted odds ratio for each of the factors included. Analyses also were stratified by the following factors individually: crash type (single vehicle, multiple vehicle), day/night, route (Interstate 5 or 90), and roadway alignment.

The researchers found that the risk of crashes was higher for double trailer trucks and single units pulling trailers, and it was also higher for younger drivers, longer hours of driving, and operation of empty trucks. The fact that long hours of driving raised the crash risk is of course of interest in the present context. Two criticisms of this study are (1) that data from drivers' logbooks on the number of hours driven per day may not be of sufficiently high quality, and (2) such analyses rest on the assumption that any characteristics not measured and not equally distributed in the crash and control populations are themselves not causal or correlated with causal factors.

A related study of the same data by the same two researchers (Jones and Stein, 1989), found that driving in excess of 8 hours compared with driving 2 hours or less resulted in a 1.8 times higher risk of a crash. In addition, driver logbook violations raised the unadjusted odds ratio by 3.0. The authors also found that a decrease in the quality of the steering raised crash risk.

Fatigue, Alcohol, Other Drugs, and Medical Factors in Fatal-to-the-Driver Heavy Truck Crashes, Safety Study **(National Transportation Safety Board, 1990)**

As mentioned above, the NTSB investigated 182 crashes in which the driver of a heavy truck was fatally injured, occurring in eight states during October 1, 1987, to September 30, 1988. The purpose of this study was to identify the principal reason(s) why each crash occurred. The great majority were single-vehicle crashes in which only the truck driver died. The sample represented about 25 percent of the crashes of this type nationally during this period. To assess the contributing factors, the NTSB developed information to augment the crash reports of police crash investigators, making it possible to describe more completely the operator(s), vehicle(s), and roadway at the time of the crashes. In addition, the NTSB interviewed representatives of the carrier, available witnesses, and reachable family members to obtain more detailed information on hours of service, fatigue, carrier operations and maintenance, safety programs, training and testing, preemployment screening, and other factors. The researchers also received the case files developed by the crash investigators and driver logs, as well as blood and urine specimens from the fatally injured drivers obtained from local coroners and medical examiners.

While at the time, the country and the NTSB were still focused on alcohol-related driving fatalities on the nation's roadways, the somewhat surprising finding of this study was that so many of these truck driver fatalities were more likely attributable to the influence of driver fatigue. The major findings relevant to driver fatigue and exceeding HOS regulations were as follows: (1) fatigue was cited as a probable cause 31 percent of the time, which made it the most frequently cited cause; (2) of the 57 drivers who were fatigued, 19 were also impaired by alcohol or other drugs; and (3) there was a strong association between HOS violations and drug use.

Commercial Motor Vehicle Driver Fatigue and Alertness Study: Project Report
(Wylie et al., 1996; Mitler et al., 1997)

The primary goal of the Driver Fatigue and Alertness Study (DFAS) was "to observe and measure the development and progression of driver fatigue and loss of alertness, and to develop countermeasures to address it, through a field study...." Beginning in 1993, 80 truck drivers aged 25-65 with at least 1 year of experience in the United States and Canada driving long-haul less-than-truckload cargo in tractor-semi-trailers were monitored for 16 weeks each as part of a naturalistic driving study. Data were collected on work-related factors "thought to influence the development of fatigue, loss of alertness, and degraded driving performance in commercial motor vehicle drivers." As is typical of naturalistic driving studies (see Chapter 5), the DFAS was carried out within an operational setting of real-life, revenue-generating trips. The work-related factors examined included the amount of time spent driving during a work period, the number of consecutive days of driving, the time of day when driving took place, and schedule regularity. For each driver, data collection lasted 4-5 days.

The drivers were divided into four groups of 20, and each group was asked to follow one of the following schedules: (1) 10 driving hours turnaround route, starting around 10 AM for five consecutive trips; (2) 10 driving hours turnaround route starting 3 hours earlier each successive day, with more night driving time than in schedule (1); (3) 13 driving hours turnaround route starting late each evening for four consecutive trips, with more night driving than in schedule (2); and (4) 13 driving hours starting in the late morning and early afternoon for four consecutive trips. Schedule (1) provided 11 hours off between trips, while the other three schedules provided only 8 hours off between trips. Measures collected for each subject included lane tracking, steering wheel movement, driving speed, distance monitoring, response vigilance tests, continuous video monitoring of the driver's face and the road ahead, and polysomnography during sleep and while driving.

The strongest factor influencing driver fatigue was determined to be time of day of driving. Drowsiness was greatest during night driving. Hours of daytime driving was not a strong predictor of observed fatigue. (Hours of nighttime driving could not be assessed as a predictor of fatigue given the study design.) Finally, there was some evidence of cumulative fatigue across days of driving. The fact that there was more than one difference among the schedules confounded attempts to interpret comparisons of means across the four groups.

Effects of Sleep Schedules on Commercial Motor Vehicle Driver Performance (Balkin et al., 2000): Study 1, Actigraphic Assessment of Sleep of CMV Drivers over 20 Days

In this field study (the report also describes a simulator study), wrist actigraphy was used to determine the hours of sleep obtained by 25 long- and 25 short-haul CMV drivers over 20 consecutive days, both on and off duty. It was found that both long- and short-haul drivers averaged 7.5 hours of sleep per 24 hours.

Stress and Fatigue Effects of Driving Longer Combination Vehicles (Battelle-Seattle Research Center, 2000)

The authors describe this study as follows:

> Between October 31, 1994, and January 21, 1995, 24 experienced longer-combination vehicle drivers drove approximately 2700 miles

> each in specially-equipped and loaded single- and triple-trailer commercial vehicles under controlled experimental conditions. The tractors were equipped with video and digital equipment to gather data on the drivers' performance. . . . Each driver who participated in the study was assigned to a specific tractor and drove it for the entire study week, using it to alternatively pull each of the three different trailer configurations: a single 48-foot trailer, a triple-trailer combination with three 28-foot trailers and standard converter (A-dollies), and a triple-trailer combination with three 28-foot trailers and double-drawbar, self-steering converter dollies (C-dollies).

With three possible configurations, there were six orderings of trailers, and four drivers were randomly assigned to each permutation. Fatigue-related measures included lane deviation assessments. Each driving day consisted of 10 hours on duty, including 8 hours of driving.

The researchers determined that driving the triple trailers contributed to increased fatigue as measured primarily by lane departures. In addition, there was substantial heterogeneity in the outcomes, including lane departures, representing 32 percent of the variability in outcomes.

Driver Sleepiness and Risk of Serious Injury to Car Occupants: Population Based Case Control Study

(Connor et al., 2002)

This study examined data from 571 car drivers involved in crashes in the Auckland region of New Zealand between April 1998 and July 1999 in which at least one occupant was admitted to a hospital or killed. It also looked at 588 controls who were randomly selected to mimic the distribution of people driving on the region's roads during the study period. The cases were "identified by cluster sampling of drivers at 69 randomly selected sites on the road network. The day of the week, time of day, and direction of travel for each survey site were randomly assigned." The goal was to determine whether the relative risk for injury was associated with various driver characteristics, especially fatigue, through case-control methods. For each driver admitted to the hospital, interviews were conducted, often within 48 hours of the crash. For crashes that were fatal to the driver, proxies were interviewed. Questions about sleep obtained made up a small portion of the interview to disguise the intent. To employ the Stanford sleepiness scale, the researchers had respondents select one of seven statements that most closely described their alertness immediately before the crash. Controls were similarly interviewed around the time of their selection.

Confounding factors considered in the analysis included age, gender, socioeconomic status, ethnicity, alcohol consumption, use of recreational drugs, time spent driving per week, vehicle speed, average traffic speed, type of road, and how long the person had been driving on the day of the crash. The analysis involved the estimation of odds ratios using logistic regression, with the complication that the cluster sample was accommodated using the SUDAAN statistical analysis software. The change-in-estimate method was used to assess potential confounders (Greenland, 1989). The confounders that passed this test were included in the logistic regression model.

The results showed a "strong association between the level of acute driver sleepiness, as measured by the Stanford sleepiness score, and the risk of injury. . . . The two direct determinants of acute sleepiness . . . sleep deprivation and time of day, were also strongly associated with the risk of an injury crash. Drivers who reported five hours or less of sleep in the previous 24 hours were at significantly increased risk compared with those who had more than five hours."

In addition to a considerable amount of missing data, this study was subject to recall bias. It is well known, as discussed earlier, that self-reports about the amount of sleep received are of uncertain quality. Finally, this was not a study of truck drivers but of automobile drivers, and it was for travel in New Zealand.

Report to Congress on the Large Truck Crash Causation Study (Federal Motor Carrier Safety Administration, 2006; Starnes, 2006)

In the Motor Carrier Safety Improvement Act of 1999, Congress mandated "a study to determine the causes of, and contributing factors to, crashes involving commercial motor vehicles." As a result, FMCSA and NHTSA conducted a "multiyear, nationwide study of factors that contribute to truck crashes. . . . A nationally representative sample of large truck fatal and injury crashes was investigated during 2001 to 2003 at 24 sites in 17 states. Each crash involved at least one large truck and resulted in at least one fatality or injury. Data were collected on up to 1,000 elements in each crash. The total sample involved 967 crashes, which involved 1,127 large trucks, 959 nontruck motor vehicles, and included a total of 251 fatalities, and 1,408 injuries."

Data collection was carried out at each crash site by a two-person team consisting of a trained researcher and a state truck inspector. They collected data on the crash scene, including information about the roadway and the weather; vehicle rollover, fire, jackknife, or cargo shift; problems with brakes, tires, steering, engine, or lights; driver credentials, method of payment, physical condition, fatigue (based on sleep pattern, work schedule, and recreational activities), and inattention/distraction;

trip start time, purpose, and intended length; and driver's familiarity with the route.

One and only one "critical event" was designated for each crash—the event that immediately led to the crash. Likewise, one and only one "critical reason" (the immediate reason for the critical event) was assigned to each event. In addition, crashes were coded with associated factors, which were indicated as being present but not necessarily causal, and more than one of these could be assigned to an individual crash. The findings relevant to fatigue were that shortage of sleep was given as the critical reason in 7 percent of the crashes, and partial sleep deprivation was given as an associated factor in 13 percent of the crashes.

The LTCCS had some design flaws. First, fatigue was assessed either indirectly or by self-report, so the quality of that information likely is not high. In addition, the requirement to find a critical event and the critical reason for that critical event could have biased the observers toward factors in immediate physical or temporal proximity. Even with these flaws, however, and the fact that the data are now 10 years old, the LTCCS is considered an important source of quality information on the relative likelihood of various causes of truck crashes.

Work Schedules of Long-Distance Truck Drivers Before and After 2004 Hours-of-Service Rule Change

(McCartt et al., 2008)[2]

In this study, three samples of long-distance truck drivers were interviewed face to face as they passed through roadside weigh stations on Interstate highways in Pennsylvania and Oregon immediately before and after the 2003 change in the HOS regulations, which increased the limit on daily driving from 10 to 11 hours. The first sample of responses was collected prior to the change, from November to December 2003; the second sample was collected 1 year after the change, from November to December 2004; and the third sample was collected 2 years after the change, from November to December 2005. A total of 1,921 drivers participated in one of the three groups of interviews.

To encourage participation, drivers were given an incentive payment of $10. Participation rates ranged from 88 to 98 percent. Questions were asked about work schedules, rule violations, and fatigued driving, with the differences between 2003 and 2004 responses and between 2003 and

---

[2] The rule changes at issue between 2003 and 2004 were (1) the daily minimum off-duty requirement was changed from 8 hours to 10 hours, (2) the maximum hours of driving prior to going off duty was changed from 10 to 11 hours, (3) the maximum tour of duty was 14 hours, and (4) the 34-hour restart period was initiated.

2005 responses providing the statistics to be interpreted. The authors report the following findings:

> The large majority (72-76% in 2004 and 69-70% in 2005) said that their current daily driving times were about the same as before the rule change. But in both 2004 and 2005, about one-fifth of drivers said they were driving more hours daily under the new rule. . . . In the 2004 and 2005 surveys, a sizeable percentage of drivers in both states reported they typically got more daily sleep under the new work rule than under the old rule. . . . At least 72% said the restart was part of their regular schedules. . . . The percentage of drivers interviewed in Pennsylvania who said they drove their trucks while sleepy at least once during the past week increased from 43% in 2003 to 48% in 2004 and then declined to 43% in 2005. . . . In Oregon, the percentage was who reported sleepy driving was 36% in both 2003 and 2004 and 41% in 2005. The percentage who reported dozing at the wheel of a truck on at least one occasion during the past month increased over time in each state, with the percentage difference between 2004 and 2005 being statistically significant.

Finally, compliance with the rules decreased in Pennsylvania over the 2-year time period but went up in Oregon.

There are two main criticisms of this study. First, as noted earlier, driver reports of sleepiness are not always of high quality. Second, there may have been other dynamics between 2003 and 2004 and 2005 that were not controlled for.

Analysis of Risk as a Function of Driving-Hour: Assessment of Driving-Hours 1 Through 11
(Hanowski et al., 2008)

This project was a naturalistic driving study of 98 CMV drivers (97 males, 1 female, age range of 24-60). Data collection started in May 2004 and was completed in September 2005. Study participants drove company trucks on their usual routes. Equipment, most of which was unobtrusive, was installed in 46 trucks to record the driver, the road ahead, and other data. The average number of weeks the drivers participated was 12.4. The final data set consisted of 2.3 million miles of driving data. Driving performance was assessed through the occurrence of critical incidents, which included crashes, near-crashes, and crash-relevant conflicts. In addition, for some of the analyses, only those incidents in which the driver was viewed as being at fault were included.

Given the potential for subjectivity in the assessment of near-crashes and crash-relevant conflicts, the number of critical incidents varied in each of eight analyses carried out. Also, to adjust for the differences in opportunities across driving hours, the frequency of critical incidents in any given driving hour was divided by the total opportunities for that hour.

One direct analysis of the frequency of critical incidents as a function of driving hour showed a visible spike in the relative frequency of critical incidents in the first driving hour, and this finding was consistent across the various analyses. There was no evidence of a time-on-task effect. A second analysis computed odds ratios using logistic regression models. The assumption of independence of incidents was not made for this analysis; instead, generalized estimating equations were used to account for correlations that might exist between and within drivers. This was done for all trips and conditional on trips that lasted the full 11 hours. In addition, there was evidence of a traffic-density effect.

Investigation into Motor Carrier Practices to Achieve Optimal Commercial Motor Vehicle Driver Performance: Phase I
(Von Dongen and Belenky, 2010)

The objective of this project was to determine the effectiveness of the 34-hour restart provision in the HOS regulations for CMV drivers. To this end, a sample of 27 healthy subjects were subjected to one of two protocols involving two 5-day work periods (14 hours per day), separated by a 34-hour restart period during which the driver transitioned back to a daytime wake, nighttime sleep schedule. The first group drove during the day and slept at night, while the second group drove at night and slept during the day. The primary outcome measure was the comparison between 10-minute psychomotor vigilance tests evaluated before and after the 34-hour restart. In addition, lane deviations were measured. Under both outcome measures, the 34-hour restart provision was more effective at mitigating the sleep loss for those working during the day than for those working at night.

Near Crashes as Crash Surrogates for Naturalistic Driving Studies
(Guo et al., 2010)

This analysis of the 100-car study (a large-scale naturalistic driving study described in Dingus et al. [2006]) examined whether safety-critical events are useful as crash surrogates. The 100-car study collected data on 2 million vehicle-miles and 43,000 hours of driving. Crashes were defined as any contact with an object at any speed in which kinetic energy is measurably transferred or dissipated. A near-crash was defined as “any circumstance that requires a rapid, evasive maneuver by the participant vehicle, or any other vehicle, pedestrian, cyclist, or animal, to avoid a crash." The researchers detected such maneuvers by looking at the vehicle kinematic data. The primary measure on which they relied to assess the impact of a factor on traffic safety was the odds ratio, that is, the odds of the presence of a factor for safety events divided by the odds of its presence for baseline events. The authors state that a measure, in this case safety-critical events, is to be viewed as a surrogate measure when (1) the causal mechanisms are the same or similar for crashes, and (2) there is a strong association between the frequency of surrogate and primary measures. Based on results of a variety of analytic techniques, the report suggests that safety-critical events are only somewhat effective as surrogates.

Hours of Service and Driver Fatigue: Driver Characteristic Research **(Jovanis et al., 2011)**

This study compared the effect of different driver HOS regulations on the odds of a crash. The analysis was based on crashes reported by the trucking companies that cooperated with the researchers involving either a fatality, an injury requiring medical treatment away from the scene of the crash, or a towaway. Driver logs for periods of 1-2 weeks prior to the crash were compared with those for two noncrash-involved drivers that were randomly selected from the same company, terminal, and month, using case-control logistic regression. Other covariates included cumulative hours driving, driving patterns over multiple days, time of day, breaks during driving, and use of the 34-hour restart policy. Data from 2004-2005 and 2010 were collected from a total of 1,564 truck drivers. Separate analyses were carried out for truckload and less-than-truckload freight modes of operation. The main results were as follows:

- Driving time was a statistically significant predictor of crash risk for the less–than-truckload drivers.
- Less-than-truckload data showed a pattern of increasing crash odds as driving time increased, with a consistent increase from hour 5 through hour 11.
- Truckload data showed significant interactions between some multiday driving patterns and increased crash risk between the seventh and eleventh hours.
- Driving breaks reduced crash risk for both types of drivers.
- Driving times that would have been a violation of the 34-hour restart provision—but were not since they occurred in 2010—were associated with an increased risk of a crash.

Issues with this study include the unclear validity of driver logs and the omission from the logistic regression model of other covariates that are predictive of crash frequency, such as driver age and experience.

An Assessment of Driver Drowsiness, Distraction, and Performance in a Naturalistic Setting
(Barr et al., 2011)

Barr and colleagues (2011) conducted a reanalysis of a naturalistic driving study carried out by Hanowski and colleagues (2000) on local and short-haul truck drivers to determine causes of fatigued driving. A major difference was that in the original analysis, attention was given only to safety-critical events. In this study, the researchers identified incidents of driver fatigue or drowsiness that occurred during all periods of driving.

The data examined consisted of 908 hours of footage (from five video cameras) on 41 drivers. The researchers reviewed the entire video library to code 3-minute durations defined either as baseline periods or as periods corresponding to the occurrence of a characteristic drowsiness behavior. The coding comprised "not drowsy, slightly drowsy, moderately drowsy, very drowsy, or extremely drowsy," and was based on evidence of yawning, rubbing eyes, closing eyes, slow blinks, bobbling one's head, and verbal announcement of drowsiness. The analysis encompassed 1,000 such baseline and fatigue events. Baseline events were matched to fatigue events using road, weather, time of day, traffic, and other conditions and served as a control group in the statistical analysis. The potential predictors available included years of driving experience, time of day of event, amount of time on duty, number of hours slept the previous night, actual sleep as measured by actigraphy, traffic density, number of lanes, the lane the driver was in, road type, road geometry, road conditions, weather, visibility, and illumination outside the vehicle.

The analysis was stratified by the coded degree of fatigue. Logistic regression models (and some contingency table and regression models) were used to determine which factors explained the difference between the baseline and fatigued events. The findings were as follows:

- Higher levels of fatigue were associated with younger drivers.
- Drowsiness was twice as likely to occur between 6:00 and 9:00 AM.
- Drivers were affected by fatigue and drowsiness in the early morning and near the end of their shift.
- Thirty percent of all severe drowsy events occurred in the first hour of the work shift.
- The relationship between sleep quality or quantity and driver fatigue was fairly weak.

- Undivided highways and poor visibility increase driver attention and reduce fatigue.

Besides the small sample size of volunteers, an important limitation of this study was its focus on local/short-haul truck drivers, who face different challenges from those faced by long-haul drivers. In particular, short-haul drivers typically do not drive at night, so the analysis could not address the impact of nighttime driving.

The Impact of Driving, Non-Driving Work, and Rest Breaks on Driving Performance in Commercial Motor Vehicle Operations

(Blanco et al., 2011)

This naturalistic driving study took place between November 2005 and March 2007. The study included 97 drivers (aged 21-73) with an average of 9 years of experience driving commercial motor vehicles. The drivers were employees of four for-hire trucking companies, and represented both long-haul operations and drivers that returned home most nights.

The drivers drove instrumented trucks during their usual routes for 4 weeks. Nine trucks were fitted with video cameras trained on the driver's face, on the steering wheel, and outside of the truck, and additional sensors measured other aspects of the driver's actions. Driver performance was assessed through the occurrence of safety-critical events, which were defined as crashes, near-crashes, and crash-relevant conflicts, as well as unintentional lane deviations. In addition, each driver was asked to fill out a daily register that included on-duty and off-duty activities for the 4 weeks of the study.

The results, based on descriptive statistics, odds ratios, and negative binomial regression models, were as follows. First, drivers spent 66 percent of their workday driving and 23 percent doing paperwork, loading or unloading, or performing other work activities. Second, using driving hour as a continuous variable in a mixed-effects negative binomial regression model to model the number of safety-critical events, Blanco and colleagues (2011) showed a statistically significant effect for time on task. Looking at the effects of individual driving hours pairwise, some of this effect stemmed from the fact that the eleventh hour had significantly more safety-critical events than the first and second hours, although there was no statistically significant difference between the effects for the tenth and eleventh hours. On the other hand, an analysis that involved counting only whether a safety-critical event had or had not occurred found no significant differences as a function of driving hour, although an analysis of the rate of occurrence of such events as a function of shift duration did show a significant increase as shifts grew longer, an effect that lasted into the fourteenth work hour. In addition, when nondriving activities were introduced during the driver's shift, such breaks significantly reduced the risk of safety-critical events for the first hour after the break. Finally, it should be noted that the spike observed during the first hour of driving in Hanowski et al. (2008) was not seen in this study.

Motorcoach Driver Fatigue Study, 2011
(Belenky et al., 2012)

This study examined whether commercial motorcoach drivers were working within the limits set by the HOS regulations. Data were collected on duty start times, total duty time per 24 hours, and total sleep time per 24 hours. Driver performance and degree of fatigue were measured and related to those predictors.

Eighty-four motorcoach drivers working for charter, tour, regular route, or commuter express carriers were studied for 31 consecutive days. During this time, they kept to their normal work/rest schedules. All drivers maintained a duty/sleep diary. Actigraphy devices were used to measure sleep/wake times, so that sleep variables could be assessed. To measure performance, drivers were administered a psychomotor vigilance test when going on and off duty and before and after breaks during the day. The Samn-Perelli Fatigue Scale and the Karolinska Sleepiness Scale were used to measure subjective sleepiness when the drivers were going on and off duty and before and after breaks. Unfortunately, baseline measures of fatigue were not available.

It was found that drivers drove an average of 43 hours per week. Total time on duty per day averaged slightly more than 9 hours and rarely exceeded the regulatory limit of 15 hours. Mean actigraphically collected total sleep duration was 7 to 9 hours. In general, drivers worked within the limits of the current HOS regulations and balanced the demands of work and family to obtain sufficient sleep.

Effect of Circadian Rhythms and Driving Duration on Fatigue Level and Driving Performance of Professional Drivers
(Zhang et al., 2014)

This naturalistic study examined the interplay among circadian rhythms, drive time, drive duration, fatigue, and driver performance. Fifteen middle-aged professional daytime drivers (using private vehicles with automatic transmission) were randomized to driving one of three schedules: (1) one group started driving at 9 AM, (2) a second group started at noon, and (3) a third group started at 11 PM. Each group drove for 6 hours. Drivers reported their rating on the Karolinska Sleepiness Scale every 5 minutes while driving so that fatigue could be assessed. Measurements of steering and lane position were taken regularly as well.

The results were as follows: (1) both circadian rhythms and increasing drive duration (between 0 and 6 hours) had significant effects on fatigue levels, and fatigue levels increased more rapidly in the evening group; (2) drivers were most likely to feel tired between 2 and 4 AM and between 2 and 4 PM; and (3) the group that was most tired was the evening group.

Asleep at the Wheel: The Prevalence and Impact of Drowsy Driving
(Tefft, 2010)
Prevalence of Motor Vehicle Crashes Involving Drowsy Drivers, United States, 2009-2013
(Tefft, 2014)

NHTSA's National Automotive Sampling System Crashworthiness Data System (NASS CDS) is a nationally representative sample of police-reported crashes. The first of these studies (Tefft, 2010) used NASS CDS data for 1999 to 2008 on crashes involving passenger vehicles that were towed from the crash scene, which represented 80,821 vehicles involved in 47,597 crashes. The researchers found that "3.9 percent of all those crashes, 7.7 percent of those that resulted in at least one person being admitted to a hospital, and 3.6 percent of those that resulted in death involved a driver who was coded as drowsy. However, the attention status of 45 percent of the drivers in the data was unknown." To address this degree of uncertainty, imputation was used to estimate the drowsy status of the remaining drivers. The result was "an estimated 7.0 percent of all crashes in which a passenger vehicle was towed, 13.1 percent of crashes that resulted in a person being admitted to a hospital, and 16.5 percent of fatal crashes involved a drowsy driver." Drowsiness was determined on the basis of information from "interviews conducted by NASS CDS investigators with crash-involved occupants from police reports." The imputation used the following covariates: maximum injury, driver injury severity, number of vehicles in crash, pre-event maneuver, crash type, day of week, hour of day, traffic flow, number of passengers, driver age, driver gender, light condition, relation to intersection, roadway alignment, speed limit, number of lanes, surface conditions, precrash critical event, vehicle disposition, year, stratum, and primary sampling unit. Among crashes in which the driver was fatally injured, information on attention was missing for 92 percent. Therefore, for inferences about crashes fatal to the driver, 92 percent of fatigue status was imputed.

This study has the following weaknesses. First, it involved only passenger vehicles (a weakness only for drawing inferences about CMV crashes). Second, the validity of assessing drowsiness from interviews is unclear. Finally, the imputation model was not validated, and it was extensively employed.

The second of these studies (Tefft, 2014) is an update of the previous study using data for 2009 to 2013. Instead of the above percentages related to crash severity of 7.0, 13.1, and 16.5 percent, the respective percentages in this later study were 6.0, 13, and 21 percent.

Summary

Table 7-1 summarizes the above studies.

HOS regulations need to take into account the trade-off between the economic advantages of transporting goods more quickly and the disadvantages of increasing crash risk. Given that, it would be helpful to make this trade-off as explicit as possible by linking increases in crash risk to increases in the number of hours of service permitted. However, given the multivariate causal structure of crashes, such a construct can be provided only by fixing all the other causal factors at levels that rarely if ever obtain, and therefore, such a statement would not be useful to support the development of policies.

Two Relevant Meta-Analyses Concerning Obstructive Sleep Apnea as a Fatigue-Related Risk Factor

Two important recent meta-analyses provide useful summaries of the literature on how obstructive sleep apnea (OSA) affects crash risk.

Tregear and colleagues (2009b) performed a meta-analysis to understand the degree to which CMV drivers with OSA are at an increased risk of crashes compared with drivers without OSA. They selected 18 studies satisfying a variety of criteria for analysis. They found that drivers with OSA had a 1.21 to 4.89 times higher crash risk compared with drivers without OSA. In addition, the 18 studies indicated that factors associated with increased crash risk for drivers with OSA are body mass index, apnea/hypopnea index, and oxygen saturation.

Tregear and colleagues (2009a) performed a meta-analysis of the impact of the use of continuous positive airway pressure (CPAP) treatment on motor vehicle crash risk for automobile and CMV drivers with OSA. Nine studies met their criteria. They found that CPAP use significantly reduced crash risk following treatment, with a 95 percent confidence interval of 0.22, 0.35.

TABLE 7-1 Summary of Studies on Hours of Service, Fatigue, and Crash Risk

Study Author(s)	Data Source/Data Collection Method	Corresponding Years	Dependent Variable
Stein and Jones (1988)	Crash reports from Washington State and random inspections	1984-1986	Crash rate
National Traffic Safety Board (1990)	Crash reports in 8 states augmented by investigators	1987-1988	Fatal-to-the-driver crash rate
Wylie et al. (1996)	Naturalistic driving study	1993	Driver fatigue as measured by lane tracking, etc.; EEG; PVT
Balkin et al. (2000)	Actigraphy	1997	Potential fatigue as measured by actigraphy; also included simulator study
Battelle-Seattle Research Center (2000)	Naturalistic driving study	1994-1995	Lane deviations
Connor et al. (2002)	Crash investigations, interviews	1998-1999	Driver ratings of sleepiness; risk of serious crash
Federal Motor Carrier Safety Administration (2006), Starnes (2006)	Crash investigations, interviews	2001-2003	Fatal truck crash data, critical event for each crash
McCartt et al. (2008)	Personal interviews	2003-2005	Interview data on fatigued driving, sleep obtained
Hanowski et al. (2008)	Naturalistic driving study	2004-2005	Frequency of safety-critical events

Analyses	Findings	Caveats
Logistic regression case control	Driving in excess of 8 hours results in a 1.8 times greater crash risk	Driver fatigue measured indirectly by time on task through logbook data
Frequencies	Fatigue cited as cause in 31 percent of sample	Driver fatigue assessed by investigators' reconstruction
Means, frequencies	Time of day important; drowsiness greatest during night driving	Confounded design
Analysis of variance	Both long- and short-haul drivers often get 7.5 hours of sleep per 24 hours	Some driver populations not subject to substantial sleep loss
Analysis of variance	Driving triple trailers adds to fatigue	Small sample; many possible confounders omitted
Logistic regression case control	Strong association between sleepiness and crash rate	Much missing data; driver recall bias re sleep obtained
Frequencies	7 to 13% of crashes associated with sleep shortage	Driver fatigue is assessed indirectly as critical events at crash sites
Frequencies, means	Drivers reported more sleep obtained under new HOS regulations; dozing increased with change in HOS regulations	Interview self-reporting on fatigue
Frequencies, logistic regression with generalized estimating equations	First driving hour is most risky; also time-of-day effect and traffic density effect	Some safety-critical events may not be evidence of fatigue

continued

TABLE 7-1 Continued

Study Author(s)	Data Source/Data Collection Method	Corresponding Years	Dependent Variable
Van Dongen and Belenky (2010)	Naturalistic driving	2008-2009	Lane deviations, PVT
Guo et al. (2010)	Naturalistic driving study	2008	Safety-critical events and crashes in cars
Jovanis et al. (2011)	Company crash data	2004-2005, 2010	Crash risk
Barr et al. (2011)	Naturalistic driving study	2000	Degree of fatigue assessed by observation
Blanco et al. (2011)	Naturalistic driving study	2005-2007	Safety-critical events
Belenky et al. (2012)	Naturalistic driving study (quasi)	2011	PVT, total time on duty, total time asleep
Zhang et al. (2012)	Naturalistic driving study, random allocation into treatments	2011 (China)	Karolinska Sleepiness Scale, lane position, steering behavior
Tefft (2010, 2014)	Automobile crash data	1999-2008	Police reports of crash frequency in automobiles

NOTES: EEG = electroencephalogram; HOS = hours-of-service; PVT = psychomotor vigilance test.

Analyses	Findings	Caveats
Frequencies, means	34-hour restart, day work shift, nighttime sleep effective at mitigating sleep loss, but work at night, sleep in day not so effective	Small sample size; many potential confounding factors
Odds ratios and Poisson regression	Safety-critical events sometimes are effective surrogates, sometimes not	Combining safety-critical events of different types may complicate inference
Logistic regression case control	Driving time was a significant predictor of crash risk for less-than-truckload trucks	Paper logs are of uncertain quality; no treatment of several confounding factors
Reanalysis of previous study; logistic regression case control	Fatigue associated with young drivers, driving between 6 and 9 AM, and when starting out	Small sample size; focus on local/short-haul truck drivers
Frequencies, odds ratios, negative binomial regression	Rate of safety-critical events increases with time on task; breaks are beneficial	Small sample size; safety-critical events are uncertain surrogates
Means	Motorcoach drivers function well under current HOS regulations	Mixed types of motorcoach drivers; no accounting for some confounding factors
Means	Fatigue is affected by circadian rhythms and by drive duration	Small sample size
Frequencies	16.5% of fatal crashes involved a drowsy driver	Assessment of fatigue by interview; much missing data

RESEARCH NEEDS

Chapter 10 suggests methodological directions for research on the relationship between fatigue among CMV drivers and highway safety. Foreshadowing that discussion, the following is a list of deficiencies that need to be addressed, some methodological and some substantive (and some not covered in Chapter 10):

- Either more experimental control of important variables is necessary for various confounding factors, or these factors need to be addressed after data collection using techniques such as propensity scoring.
- More research is needed on the relationship among HOS regulations, driver fatigue, and crash risk for bus drivers.
- More research is needed to disentangle the effects of different causes of fatigue—sleep deprivation; chronic lack of sleep; circadian displacement; times on task; and medical conditions, including OSA.
- More research is needed on how drivers decide how to address their sleepiness while driving, that is, what do they do while driving after determining that they are fatigued.
- More research is needed on the benefits of CPAP as treatment for CMV drivers with OSA; perhaps safe studies might be conducted in driving simulators.

8

Fatigue and Health and Wellness

This chapter begins by describing what is known about the linkages between fatigue and health. Following a brief summary of the medical certification of the health of commercial motor vehicle (CMV) drivers, the discussion turns to a medical condition that has been the subject of a great deal of attention with respect to the health of CMV drivers and the implications for highway safety—obstructive sleep apnea (OSA). The discussion addresses OSA among CMV drivers, the association of severe OSA with increased crash risk, treatment of OSA with positive airway pressure (PAP) devices, and medical examination policy for drivers regarding OSA. Next, the chapter reviews non-OSA medical conditions among CMV drivers and the linkages between lifestyle factors and drivers' health. The final section reviews current fatigue and health and wellness management programs for CMV drivers.

LINKAGES BETWEEN FATIGUE AND HEALTH[1]

As mentioned in the statement of task for this study (see Box 1-1 in Chapter 1), the panel was to "assess the relationship of these factors [hours of driving, hours on duty, and periods of rest] to drivers' health over the longer term."

---

[1] Here, as is true throughout this report, the term *fatigue* is used although the discussion in Chapter 3 supports focusing on insufficient sleep, a primary component of fatigue.

A substantial evidence base supports the fundamental relationship between sleep needs and health risks. According to the report *Sleep Disorders and Sleep Deprivation: An Unmet Public Health Problem* (Institute of Medicine, 2006), "The cumulative long-term effects of sleep loss and sleep disorders have been associated with a wide range of deleterious health consequences including an increased risk of hypertension, diabetes, obesity, depression, heart attack, and stroke. After decades of research, the case can be confidently made that sleep loss and sleep disorders have profound and widespread effects on human health." In its *Mortality and Morbidity Weekly Report* of March 4, 2011 (Buxton et al., 2010; Strine and Chapman, 2005), the Centers for Disease Control and Prevention (CDC) states, "Sleep impairment is linked as a contributing factor to motor vehicle crashes, industrial disasters, and medical and other occupational errors. Persons experiencing sleep insufficiency are more likely to have chronic diseases such as cardiovascular disease, diabetes, depression, or obesity." Further, a recent survey by Czeisler (2015) summarizes the linkages between fatigue and many of the health problems exhibited by truck and bus drivers:

> . . . in the 15 years since Eve Van Cauter and her colleagues at the University of Chicago discovered that sleep deficiency adversely impacts metabolic and endocrine functions, it has been demonstrated unequivocally that the duration, timing, and quality of sleep also critically affect physical health, mental health, performance, and safety. Thus it is clear that sleep is critical not just for optimal brain functioning but also for optimal functioning of the body as well. . . .
>
> Rigorous physiological studies have demonstrated that just a week or two of sleep curtailment increases appetite and food intake, decreases insulin sensitivity and glucose tolerance, . . . degrades the ability to resist infection, disturbs mood, increases the vulnerability to attentional failures, and when combined with prior chronic circadian disruption, impairs pancreatic β-cell responsiveness. Concurrently, epidemiologic studies have revealed that habitually short sleepers have an increased prevalence of obesity; . . . that habitually short and habitually long sleepers are at increased risk for incident calcification of the coronary arteries, incident coronary heart disease, incident type 2 diabetes, incident stroke, and death. . . . During the same time interval, emerging evidence that chronic exposure to recurrent disruption of sleep and circadian timing induced by night shift work increases the risk of breast cancer, endometrial cancer, colorectal cancer, and prostate cancer. . . . Moreover, extended duration (> 24 hours) work shifts have been associated with poorer performance on clinical tasks, and increased risks of serious medical errors, preventable adverse events, self-inflicted percutaneous injuries, and motor vehicle crashes among resident physicians.

> . . . Epidemiologic studies reveal that night shift work is associated with increased odds of obesity; a 5-fold increase in the risk of progressing from impaired glucose tolerance to diabetes; an increased risk of blood pressure elevation; incident hypertension; incident coronary heart disease, including fatal and nonfatal myocardial infarctions; and that a decade of exposure to shift work chronically impairs cognition.

It is important to be clear that the causal arrow can go in both directions, with fatigue raising the risks of many health problems and some health problems raising the risks of fatigue. To take an important example, as Czeisler (2015) argues, fatigue may increase the risk of obesity, and obesity can increase the potential for OSA, which can cause fatigue. Further, it is obvious that some obesity is not caused by fatigue, and some fatigue is not caused by obesity. For example, the unhealthy work conditions of many truck driving situations can lead to adverse health consequences, including obesity, which can occur regardless of whether drivers obtain sufficient sleep. More generally, the precise degree to which a shortage of sleep results in poor health or poor health leads to a shortage of sleep is unknown. The key point is that the two have an important interaction. This interaction warrants examining what efforts are being made to educate truck and bus drivers about these linkages and to motivate them to make changes that can reduce the incidence of these various health problems. Further, it is of particular interest in the present context to consider how such interventions can be evaluated.

> *Conclusion 1: Insufficient sleep can increase the risk of developing various health problems, including obesity, diabetes, hypertension, and cardiovascular disease, all of which can impact an operator's level of alertness while driving and potentially impact crash risk.*

MEDICAL CERTIFICATION OF THE HEALTH OF CMV DRIVERS

An unhealthy driver behind the wheel of a commercial motor vehicle compromises the safety of the driver and general public. To address this issue, the Federal Motor Carrier Safety Administration (FMCSA) requires that CMV drivers operating in interstate commerce maintain a current medical examiner's certificate to drive. CMV drivers must be examined at least every 2 years by a certified medical examiner to ensure that they are fit to operate their vehicle without risk of sudden or gradual impairment or incapacitation. Following this medical exam, the medical examiner can certify the driver for up to 2 years, disqualify the driver, or impose various intermediate actions.

Medical examiners follow a set of 13 federal medical standards when conducting the medical exam.[2] Of these 13 standards (which have been in effect, with only minor changes, since 1971), the following 4 result in suspension of the commercial driver's license (CDL) unless the driver has been granted an exemption by FMCSA: (1) insulin-requiring diabetes, (2) seizures requiring the use of antiseizure medication, (3) vision requirements, and (4) hearing requirements. The certification determinations for the other 9 standards are left to the discretion of the medical examiner. Certification determinations for certain medical conditions also are left to the discretion of the medical examiner, with FMCSA offering guidance on many such conditions. The difference between medical standards/regulations and medical guidelines is that medical standards must be met by CMV drivers and are to be verified by the medical examiner, whereas medical guidelines (including advisory criteria and medical conference reports) are suggestions for best practices to be utilized by the medical examiners. Medical standards and medical guidance issued by FMCSA are included in the *Medical Examiner Handbook*, the primary reference tool for medical examiners (see the discussion of this and other resources in the next section).[3]

OBSTRUCTIVE SLEEP APNEA

OSA is a health condition in which the airway becomes partially blocked during sleep, resulting in frequent awakenings. As a result, the person experiences daytime sleepiness. OSA is linked to other medical conditions, such as diabetes and various cardiovascular diseases (see Surani, 2014). Its severity is judged using the apnea-hypopnea index (AHI), which measures the number of awakenings or hypopneas one experiences per hour of sleep (Ruehland et al., 2009). According to Snyder (2013), "Mild OSA is present when the AHI is between 5 and 15 and OSA symptoms exist. This means the person has episodes of delayed breathing five to fifteen times in an hour. Many people with AHIs in this range have no symptoms at all. Moderate OSA is defined as an AHI between 15 and 30,

---

[2] The 13 standards can be found at https://www.fmcsa.dot.gov/regulations/title49/section/391.41 [March 2016].

[3] Until recently, medical examiners had the FMCSA-issued *Medical Examiner Handbook*, as well as Frequently Asked Questions and interpretations of the regulations. The *Medical Examiner Handbook* was removed from the FMCSA website in early 2015 for update, although most examiners still use it as a resource. Examiners can also refer to recommendations of the Medical Expert Panel on Obstructive Sleep Apnea and the Medical Review Board (MRB), although these recommendations have not been adopted or accepted by FMCSA. The MRB, established in 2006 by FMCSA, comprised several practicing physicians, chosen from a field of many qualified candidates, with a wide variety of expertise and experience.

regardless of the presence of symptoms, while an AHI greater than 30 is termed severe OSA." It should be noted that these categorizations are arbitrary and are unrelated to a specific degree of performance impairment.

Pack and colleagues (2002) measured a random sample of CDL holders within 50 miles of the University of Pennsylvania and found that 17.6 percent had mild OSA, 5.8 percent had moderate OSA, and 4.7 percent had severe OSA. A study by Berger and colleagues (2012) involved administering a computer-based screening instrument to 19,371 drivers across three trucking firms, 5,908 (30%) of whom were classified as at high risk for OSA. Of those, a random 2,103 underwent polysomnography, and 68 percent of them had an AHI greater than 10. This finding suggests that at least 20 percent of CMV drivers (.68 × 5,908/19,371) have at least mild OSA. What is more difficult to establish is the extent to which varying degrees of OSA are associated with different degrees of impaired performance for CMV drivers.

Increased Risk of Crashes for Commercial Motor Vehicle Drivers with OSA

A substantial research literature supports the conclusion that severe OSA—that is, OSA with an AHI over 30—is associated with increased crash risk for the nonprofessional driver. Some of the major contributors to this literature include Aldrich (1989), Findley et al. (2000), George and Smiley (1999), Teran-Santos et al. (1999), and Wu and Yan-Go (1996). (See Smolensky et al. [2011] for a summary of studies analyzing the relationship between OSA and frequency of crashes.) The reason for this increased crash risk is that OSA causes cognitive dysfunction, due largely to sleep fragmentation (Bedard et al., 1993; Day et al., 1999; Deaconson et al., 1988; Drummond et al., 2000; Feuerstein et al., 1997; Kaneko et al., 2003; Kim et al., 1997).

Besides fatigue, the cumulative effects on the brain of chronic repeated nocturnal hypoxic episodes may cause irreversible cognitive deficits (Bedard et al., 1993; Nowak et al., 2006). Evidence indicates that some individuals with OSA are unaware of their cognitive impairment or even of being drowsy (Dement et al., 1978; Engleman et al., 1997). Several studies have shown little or no correlation between an OSA patient's perception of sleepiness and motor vehicle crash history (Barbe et al., 1998; Horstmann et al., 2000; Teran-Santos et al., 1999; Yamamoto et al., 2000; Young et al., 1997). Because of this lack of insight, these individuals are less likely to restrict themselves from driving than drivers who are more aware of being drowsy, despite being at increased risk for a crash.

Tregear and colleagues (2009b) performed a systematic review and meta-analysis of 18 studies drawn from seven databases described in those sources. Two of these 18 studies examine the association between OSA and crash risk for CMV drivers, while the rest are focused on passenger car drivers. Nine of the studies provide data on the relative incidence of crashes among comparable individuals with and without the disorder. Pooling of these data revealed that the mean crash risk ratio associated with OSA was likely to fall within the range of 1.30 to 5.72. Thus, if the underlying crash risk for a driver is 0.08 crashes per person-year, the crash risk for a driver with OSA can be expected to be in the range of 0.10 to 0.46 crashes per person-year. A series of sensitivity analyses was used to establish the robustness of these results to various assumptions. While the quality of the individual studies is not high, the data are qualitatively consistent, making it unlikely that future studies will contradict the finding that nonprofessional drivers with OSA are at increased risk for a motor vehicle crash. The strength of this conclusion results from the different environments and situations tested and methods used.

The question then is whether there also is an increased risk for crashes among CMV drivers with severe OSA (and what that increased risk is). A finding regarding CMV drivers reported by Tregear and colleagues (2009b) is that as "a group, drivers with OSA are at an increased risk for a motor vehicle crash when compared with comparable drivers who do not have the disorder even though a precise estimate of the magnitude of this increased risk could not be determined." An argument might be made that definitive proof of an elevated crash risk among CMV drivers with OSA of a given hypopnea level can come only from a randomized controlled trial examining whether untreated OSA is associated with a higher risk of crashing. Such a study, however, is not feasible; it would also be unethical. Once professional drivers have been diagnosed with OSA, they must initiate treatment by using a PAP device, or they cannot continue driving. Therefore, professional drivers with OSA cannot legally or ethically be tested on the open road unless their OSA is addressed. (Such studies could be conducted in driving simulators with, say, a 2-week confinement, in which the drivers' sleep environment and schedule were controlled; see O'Neill et al. [1999].) The same barriers would forbid the testing of CMV drivers with OSA to assess the benefits of PAP devices in reducing crash risk.

There are reasons to believe that nonprofessional and CMV drivers would differ in their response to OSA. They certainly are different populations of drivers and faced with different tasks, so the hypothesis that CMV drivers with OSA are not at increased risk needs to be considered seriously. Professional drivers may take more naps when needed, they may orient their schedules to allow for more rest, they may ingest greater amounts of caffeine, or they may avoid congested areas when they feel tired. On the other hand, CMV drivers often drive longer hours and often are older than the average noncommercial driver studied. Evidence also suggests that, relative to noncommercial drivers, they suffer from more ailments that could increase their degree of fatigue. Evidence that could be used to reject the notion that there is a difference between the two categories of drivers would support the hypothesis that the findings for nonprofessional drivers of passenger cars can be transferred to professional drivers of trucks and buses. Conversely, evidence that professional drivers take actions such as those cited above more often or more effectively relative to nonprofessional drivers would support the notion that there is no increased risk of crashes among the former group.[4]

The panel did not find credible evidence that actions to counter fatigue are taken more often or more effectively by CMV drivers. However, an unpublished study by Schneider National found that the crash frequency among drivers using PAP devices was reduced compared with their previous crash record.[5] In addition, literature on somewhat analogous populations, such as railroad engineers, those in charge of ships, and airline pilots, provides evidence that OSA is a factor in increased risk of mistakes and that use of PAP devices reduces that risk (see, e.g., Quan and Barger, 2015).

Finally, it can be argued that the absence of proof is not proof of the converse. Unless good studies show that CMV drivers with severe OSA are not at increased risk of crashes, it is a principle of preventive medicine and a reasonable precaution to accept that the parallel work among general drivers applies to CMV drivers as well. Increasingly, observational science is being held to an impossible standard—either it is proven that the adverse event occurs in the "exact" population of interest, or there is considered to be no proof. The panel believes many preventable losses will be suffered if a very narrow view of what constitutes an evidence base is used to argue that only evidence that is direct and not indirect is sufficient.

The panel notes that the relationship between degree of OSA and increase in crash risk is likely represented by a smooth response curve, such that those drivers with mild OSA have little or no greater risk of crashes than drivers without OSA. Therefore, it would be extremely helpful if in the future, the research literature could be more precise about the level of OSA being considered.

---

[4] The severity of the crash is a factor that can impact findings on the association between OSA and crash risk. Stevenson and colleagues (2014) carried out a case-control study using survey questionnaire data and found that OSA was not associated with nonfatal, nonsevere crashes.

[5] Presentation by Don Osterberg, Schneider National to the Panel on Research Methodologies and Statistical Approaches to Understanding Driver Fatigue Factors in Motor Carrier Safety and Driver Health in Washington, D.C., on May 28, 2014.

> *Conclusion 2: Based on the evidence on drivers who are not commercial motor vehicle drivers, obstructive sleep apnea is known to increase crash risk, and there is no evidence base or compelling reason for thinking that the same would not also be true among commercial motor vehicle drivers.*

Treatment for OSA

The only treatment for OSA besides weight loss that has shown widespread efficacy is the use of positive airway pressure (PAP) devices. For nonprofessional drivers with OSA, use of such devices is associated with a reduced risk of crashes compared with those with OSA who do not use such devices (Cassel et al., 1996; Findley et al., 2000; George and Smiley, 2001; Tregear et al., 2009b). A key question, then, is whether the use of PAP devices also reduces the risk for CMV drivers, and if so, to what extent?

As in the case of linking OSA and crash risk, while studies linking PAP use to decreased crash risk have some limitations, the panel believes the finding of reduced risk for nonprofessional drivers can be strongly supported. However, there is not a substantial literature *directly* showing that the same assertion can be made for CMV drivers (with a given degree of hypopnea).

> *Conclusion 3: Better understanding is needed of the effects of treating obstructive sleep apnea in commercial motor vehicle drivers with positive airway pressure (PAP) therapy with respect to the amount and quality of sleep they obtain and their cognition and driver performance following PAP treatment sessions.*

A major complication is that acceptance of and adherence to PAP therapy is a problem for many patients (Kribbs et al., 1993; Weaver et al., 1997). Moreover, the amount of PAP use needed to produce clinically meaningful improvements in real-world sleep, cognition, and behavior remains unclear (Gay et al., 2006).

Medical Examination Policy Regarding OSA

FMCSA's current medical examination policy for CMV drivers states, "A person is physically qualified to drive a motor vehicle if that person has no established medical history or clinical diagnosis of a respiratory dysfunction likely to interfere with his/her ability to control and drive a motor vehicle safely." Thus if the medical examiner detects a respiratory dysfunction that in any way is likely to interfere with the driver's ability to control and drive a commercial motor vehicle safely, the driver must be referred to a specialist for further evaluation and therapy before being certified. While no regulation specifically mentions OSA, it would be covered under this respiratory standard.

In 1998, OSA was first mentioned in the advisory criteria for the respiratory standard:

> There are many conditions that interfere with oxygen exchange and may result in incapacitation, including emphysema, chronic asthma, carcinoma, tuberculosis, chronic bronchitis and sleep apnea. . . .

Medical examiners could find minimal guidance on OSA in a 1991 *Conference on Pulmonary/Respiratory Disorders and Commercial Drivers* (Turino et al., 1991), sponsored by the Office of Motor Carriers, a predecessor of FMCSA. The experts on that panel recommended that if there were any suspicion that a driver had OSA, the driver should be evaluated and the condition successfully treated before the driver returned to work. Medical examiners were offered no specific criteria for having concern or suspicion that a driver was at risk of having OSA. The conference participants recommended at least a 1-month waiting period after treatment for OSA had been initiated before a driver returned to commercial driving. They also advised evaluation of the effectiveness of treatment through either multiple sleep-latency testing (MSLT) or polysomnography. They recommended further that drivers with OSA who were medically qualified to drive commercial motor vehicles be reevaluated annually by means of sleep studies or MSLT.

The 1991 conference report was not widely disseminated, and the majority of medical examiners likely were unaware of its existence. In 2000,[6] the examination form used to evaluate CMV drivers was updated and for the first time included a specific question on OSA, asking the driver to indicate "yes" or "no" with respect to "sleep disorders, pauses in breathing while asleep, daytime sleepiness, loud snoring." When this form was first issued, many drivers responded "yes" to this question, and examiners reported that they either asked these drivers to provide documentation from their treating provider or referred the drivers for further evaluation. By the second year of use of this form, examiners indicated that few drivers checked "yes" for this question.

In preparation for the National Registry of Certified Medical Examiners (NRCME), FMCSA compiled the *Medical Examiner Handbook* from

---

[6] U.S. Department of Transportation Federal Motor Carrier Safety Administration. (2000). Final rule. Physical qualification of drivers; medical examination; certificate. *Federal Register, 65*(194), 59363-59380. Available: http://www.gpo.gov/fdsys/pkg/FR-2000-10-05/pdf/00-25337.pdf [March 2016].

existing guidance drawn from various FHWA (or FMCSA)-sponsored panels, frequently asked questions (FAQs), advisory criteria, and regulatory guidance. The section with information on OSA was the last section posted—not until fall 2010—and was not significantly different from what had been in the 1991 conference report.

The *Medical Examiner Handbook* indicated that examiners "should not certify the driver with suspected or untreated sleep apnea until etiology is confirmed and treatment has been shown to be stable, safe, and adequate/effective," but still offered no specifics on how to identify drivers suspected of having OSA. Additional guidance for examiners had been provided in three FAQs on the FMCSA website addressing sleep disorders. FMCSA also had a section of its website devoted to sleep apnea, "Spotlight on Sleep Apnea," with information for drivers and carriers, but no information on screening, diagnosis, or treatment. In early 2014, to comply with Public Law 113-45,[7] FMCSA removed the section of the *Medical Examiner Handbook* on OSA, the FAQs on sleep disorders, and "Spotlight on Sleep Apnea" from the FMCSA website. In early 2015, as noted earlier, the entire *Medical Examiner Handbook* was removed from the FMCSA website for revision.

There have been several efforts to develop criteria that examiners could consider for identifying drivers at highest risk of having OSA. In 2006, a Joint Task Force of the American College of Occupational and Environmental Medicine, the American College of Chest Physicians, and the National Sleep Foundation recommended OSA screening criteria for drivers (Hartenbaum et al., 2006). In 2007 and 2008, based in part on an FMCSA-supported evidence report (updated in 2011) (Williams et al., 2011), FMCSA's Medical Expert Panel on Obstructive Sleep Apnea and Commercial Motor Vehicle Driver Safety (MEP) (Ancoli-Israel et al., 2008) and Medical Review Board (MRB) (Federal Motor Carrier Safety Administration Medical Review Board, 2008) developed recommendations, referenced earlier, that would have required testing a larger percentage of drivers than the Joint Task Force had proposed.[8] In 2012, the Motor Carrier Safety Advisory Committee (MCSAC), an industry advisory group, and the FMCSA MRB offered joint recommendations (Federal Motor Carrier Safety Administration Medical Review Board, 2012).

FMCSA did not adopt the three sets of recommendations described in the previous paragraph. Examiners therefore continue to vary in how

---

[7] Public Law No. 113-45 Commercial Motor Vehicle Operator Requirements Relating to Sleep Disorders. Signed by President Obama October 13, 2013. Available: http://www.gpo.gov/fdsys/pkg/PLAW-113publ45/pdf/PLAW-113publ45.pdf March 2016].

[8] The need to require testing of a larger percentage of drivers than proposed by the Joint Task Force would have been linked to the establishment of more stringent criteria (e.g., lowering the body mass index [BMI] threshold for determining that a driver must be referred for sleep testing).

they evaluate drivers who may be at risk of having OSA. Durand and Kales (2009) found that while members of the American College of Occupational and Environmental Medicine believed that screening for OSA was important, fewer than 50 percent used any specific criteria for such screening.

Until 2014, medical examiners could be any health care providers licensed by their state to perform physical examinations, including medical doctors; doctors of osteopathy; nurse practitioners; physician assistants; and in some states, chiropractors, dentists, or even physical therapists. The NRCME was fully implemented in 2014,[9] and examiners on the registry are now required to be trained and certified in accordance with the NRCME Core Curriculum, which focused on the regulations, guidelines, and other official information from FMCSA available at that time.

With implementation of the NRCME, it was hoped that some consistency among medical examiners would be established, as all examiners were required to complete training that at least met the FMCSA Core Curriculum. While some training programs were limiting instruction to what was in the *Medical Examiner Handbook*, others were introducing additional resources, such as the MRB, MEP, or MCSAC recommendations, and still others were teaching (erroneously) that there were specific required criteria for screening drivers for OSA. The NRCME *Medical Examiner Sample Training Handbook* (National Registry of Certified Medical Examiners, 2012) noted that training programs could teach material beyond the *Medical Examiner Handbook*, provided they clearly highlighted that the material was not endorsed by FMCSA, and that examiners could use more current guidance than that issued by FMCSA in making certification determinations. FMCSA had indicated that, while the NRCME certification examination would include only information that had undergone public notice and comment, examiners should consider current best practice in making certification determinations, not only for OSA but also for other conditions for which official information from FMCSA was lacking, such as Parkinson's disease or use of potentially impairing medications such as opioids.

FMCSA still has not provided specific guidance for medical examiners—this despite recommendations to the agency from the National Transportation Safety Board (NTSB) (2009); the expert recommendations of the FMCSA MEP, MRB, and MCSAC; and requests from the American

---

[9] U.S. Department of Transportation Federal Motor Carrier Safety Administration. (2012). Final rule. National registry of certified medical examiners. *Federal Register, 77*(77), 24104-24135. Available: http://www.gpo.gov/fdsys/pkg/FR-2012-04-20/pdf/2012-9034.pdf [March 2016].

College of Occupational and Environmental Medicine (ACOEM).[10] The absence of specific guidance to certified medical examiners on assessing CMV drivers for OSA presents challenges for employers who are relying on the examiners to make that determination but finding that inconsistent criteria are used, even within the same examiner group. There now exists a risk of grievances or legal action against employers or examiners that attempt to utilize current best practices and require diagnostic studies for some drivers. At the same time, lawsuits have been brought against employers and examiners when a crash has occurred; and the CMV driver subsequently has been found to have OSA but was not tested even though he or she would have been considered at high risk of the disorder based on one or more of the above recommendations. In January 2015,[11] FMCSA issued a bulletin to medical examiners and training associations stating that examiners should use current best practices in determining which drivers should have objective testing and offering some considerations with respect to OSA, but noting that FMCSA has no specific requirements related to the disorder. Many examiners and employers have indicated that they will limit screening drivers for OSA until FMCSA issues specific guidance.

A 2014 survey (Hartenbaum, 2015) of examiners, both physicians and nonphysicians, both ACOEM-trained and not, who were listed on the NRCME demonstrated no consistency in examination outcomes for five scenarios in which current best practice would recommend that drivers have diagnostic testing for OSA. Selections included certify for 2 years, certify for 1 year and instruct driver to discuss risk of OSA with primary care physician, certify for 3 months and request documentation from primary care physician on risk of OSA, certify for 3 months and require sleep study, or not certify at all. The appropriate approach to screening CMV drivers for OSA remains contentious, and the absence of specific guidance affects drivers, examiners, employers, and third-party administrators.

> *Conclusion 4: It is apparent that medical examiners who certify commercial motor vehicle drivers are not consistent in understanding and applying current best practice to identify drivers who may be at risk of moderate to severe obstructive sleep apnea. Many examiners therefore*

---

[10] American College of Occupational and Environmental Medicine (ACOEM) letter to Congressmen Howard and Mica, June 28, 2010; and ACOEM letter to Administrator Ferro, July 24, 2013.

[11] FMCSA Bulletin to Medical Examiners and Training Organizations Regarding Obstructive Sleep Apnea. Available: https://nationalregistry.fmcsa.dot.gov/NRPublicUI/documents/OSA%20Bulletin%20to%20MEs%20and%20Training%20Organizations-01122015.pdf [March 2016].

are inconsistent in making determinations as to when a driver should be referred for additional sleep testing.

OTHER MEDICAL CONDITIONS

The panel has no conclusions or suggestions for change regarding the diagnosis of non-OSA medical conditions that are associated with fatigue and can result in suspension of CMV driving privileges. Nonetheless, this section summarizes the evidence that is required for diagnosis of other medical conditions based on the FMCSA *Medical Examiner Handbook*.

Hypertension

Hypertension—elevated systolic or diastolic blood pressure—is a risk factor for cardiovascular disease, peripheral vascular disease, and chronic renal insufficiency. A person with a systolic blood pressure reading of greater than 140 or a diastolic reading of greater than 90 is regarded as having hypertension. CMV drivers are more likely to have hypertension than their peers in other professions (and there is some evidence that this gap is widening) (Thiese et al., 2015a, 2015b). The salient regulation in the *Medical Examiner Handbook* is as follows: "If a driver has hypertension and/or is being medicated for hypertension, he or she should be recertified more frequently. An individual diagnosed with Stage 1 hypertension (BP is 140/90-159/99) may be certified for one year. At recertification, an individual with BP equal to or less than 140/90 may be certified for one year; however if his or her BP is greater than 140/90 but less than 160/100, a one-time certificate for 3 months can be issued. . . . Once the driver has reduced his or her BP to equal to or less than 140/90, he or she may be recertified annually thereafter."

Diabetes Mellitus

About 8 percent of the population has diabetes, and one-third of cases may be undiagnosed. By contrast, the prevalence of diabetes in CMV drivers may be as high as 30 percent (Thiese et al., 2015a, 2015b), and evidence indicates that the incidence of diabetes in CMV drivers increased between 2005 and 2012. Factors that can affect blood glucose control are fatigue, lack of sleep, poor diet, and stress. Drivers diagnosed with diabetes mellitus who require insulin to control their condition are not eligible for certification because poorly controlled blood glucose can lead to fatigue, lethargy, and sluggishness, and complications can lead to acute hyperglycemia, which in turn can cause seizures, loss of consciousness, and impaired cognitive function. The medical examination relevant to diabetes includes assessment of glycosuria, retinopathy, macular degeneration, peripheral neuropathy, coronary heart disease, cerebrovascular disease, autonomic neuropathy, and nephropathy.

Cardiovascular Problems

Thiese and colleagues (2015a, 2015b) suggest that up to 5 percent of CMV drivers suffer from cardiovascular disease. The primary manifestations include acute myocardial infarction, angina pectoris, and congestive heart failure. Drivers that have cardiovascular disease, relative to those without, are more likely to experience sudden death or incapacitation. Since other conditions may exacerbate cardiovascular problems, the decision on certification of those with cardiovascular disease depends on a comprehensive medical assessment of overall health. However, if a driver is diagnosed with myocardial infarction, angina pectoris, coronary insufficiency, thrombosis, or any other cardiovascular disease known to be accompanied by syncope, dyspnea, collapse, or congestive cardiac failure, he or she is not to be certified for driving. The diagnosis includes evaluation for heart murmurs, extra heart sounds, arrhythmias, an enlarged heart, abnormal pulse, carotid or arterial bruits, or varicose veins. Details on cardiovascular problems and the increased risk for truck drivers can be found in FMCSA (2002) and the references included therein.

LIFESTYLE FACTORS AND CMV DRIVERS' HEALTH

Long-haul CMV drivers face clear challenges to their health and wellness, including a sedentary lifestyle, limited access to healthy food, and sleep pressures. Based on a recent survey of long-haul truck drivers, Sieber and colleagues (2014) report that about 69 percent are obese (compared with 35.5% of the general population over age 20) (National Center for Health Statistics, 2015); 51 percent smoke (compared with 17.8% of adults older than 18) (National Center for Health Statistics, 2015); and few exercise regularly, which contributes to cardiac problems, diabetes, high blood pressure, and other morbidity. In addition, as discussed in Chapter 3, many commercial truck drivers get less than 6 hours of sleep at night (National Sleep Foundation, 2012), which is insufficient for most adults to maintain acceptable levels of alertness on the job and may eventually lead to adverse health consequences (Czeisler, 2015).

The panel emphasizes that these are broad generalizations that do not pertain to all CMV drivers. There are wide differences in drivers' working conditions and lifestyles, work-rest schedules, and all related factors (e.g., sleep opportunities, availability of nutritious food, exercise facilities, work stressors, family life interactions). Truly comprehensive surveys in this area have not been carried out, and given the diversity of CMV drivers, multiple surveys may be needed. In any case, much remains to be learned about fatigue, health behaviors, and their effects on CMV drivers.

> *Conclusion 5: Substantial data gaps limit understanding of factors that impact the health and wellness of commercial motor vehicle drivers. Closing these gaps would aid greatly in developing a better understanding of drivers' current status and long-term prospects with respect to health and wellness.*

That being said, the following description is generally correct. Long-haul commercial truck and bus drivers have unusual working conditions and lifestyles (see Chapter 2). Many contend with pervasive on-the-job stressors, such as the need to meet delivery schedules; constant traffic on the roadways; and monetary pressures given their income levels, which motivate them to work longer hours. Over-the-road truck driving may necessitate being away from home for weeks at a time, missing family members while away; facing extended periods of boredom while driving; sleeping in different settings every night, including use of sleeper berths; dining mainly in fast food restaurants; lacking adequate time or facilities for exercising; and remaining in sedentary work postures (sitting all day in the cab seat).[12] CMV drivers' use of tobacco products likely is greater than that among the general public (e.g., a National Institute for Occupational Safety and Health [NIOSH] survey [Sieber et al., 2014] found that 51% of the drivers surveyed smoked). As discussed in earlier chapters, drivers often must follow irregular work schedules that disrupt physiological bodily rhythms, affecting sleep and health, and they frequently experience degraded on-the-job alertness. While all CMV drivers may not have these same issues and many work more normal work schedules, they still have jobs that are primarily sedentary and pose other health risks.

There are some notable differences in driver regimens between commercial truck and bus drivers. Some bus drivers (e.g., those who drive tour buses) may work split schedules (tourists spend several hours on site before departing again for the next stop), and drivers handle passengers' luggage for them. Some bus drivers work in two-person teams, but they have no sleeper berths; while en route, the second driver must sleep in

---

[12] The association between whole-body vibration in truck driving and back problems was investigated in a NIOSH study (Blood et al., 2015). The study found that being exposed to seated whole-body vibration for extended periods is one of the leading risk factors for the development of low back disorders. Thus truck and bus drivers are at risk of developing back problems as they are regularly exposed to continuous whole-body vibration while driving heavy vehicles.

an awkward seated posture. By contrast, trucking team drivers may have comfortable sleeper berths for use by the second driver while the truck continues to move (Krueger and Van Hemel, 1997).

Partly as a result of their lifestyle, many long-haul truck drivers suffer from one or more medical conditions. As discussed above, they often have diabetes and suffer from cardiovascular problems; they may have undiagnosed sleep disorders such as OSA; they often experience musculoskeletal disorders (e.g., back pain) and injuries; and many are in generally poor health, which may adversely affect their driving performance and safety (see, e.g., Krueger, 2008; Krueger et al., 2007; Saltzman and Belzer, 2007). Moreover, many long-haul truck drivers lack access to adequate health care, especially while they are away from home. (Being constantly on the move makes it difficult to access regular health care when needed.) A NIOSH survey of U.S. long-haul truck drivers found that beyond lacking access to health care, 38.1 percent of drivers in the survey sample were not covered by health insurance or a health care plan. A small-scale cross-sectional study yielded similar findings (Apostolopoulos et al., 2010), with 33 percent of 316 participating truck drivers lacking health insurance. Not having access to adequate health care is especially common among CMV drivers, such as independent owner-operators, who do not work for large carriers. Large commercial carriers are more likely to arrange for health care and insurance programs for their employees. Many self-employed, owner-operator drivers who lack health insurance have been reported to be reluctant to obtain needed health care (Krueger, 2012; Saltzman and Belzer, 2007). Conceivably, with the implementation of the Affordable Care Act, some of these drivers will have improved access to health care.

Conclusion 6: Many commercial motor vehicle drivers work/live with occupational pressures that adversely affect health.

Given the impact of various aspects of the lifestyle of CMV driving, it is important to develop a better understanding of what aspects of that lifestyle are associated with negative health outcomes. There are many suggestions in the literature cited above and in previous chapters that CMV drivers' insufficient sleep, or poor diet, or lack of exercise, or constant jostling, or stress contributes to their health problems. There are also suggestions that CMV driver's compensation schemes contribute to their sleep insufficiency. Unfortunately, progress on isolating the factor(s) with the greatest impact on drivers' health and understanding the extent to which behavioral changes result in improvements is greatly hampered by the lack of data in two important areas: (1) how CMV drivers live (especially how much sleep they get), and (2) how behavioral changes are related to improved health outcomes. As was mentioned in Chapter 5, very few surveys of drivers have been carried out, and even in those surveys, reliable information on such important factors as the amount of sleep obtained and other aspects of the respondents' lifestyle was not collected. Furthermore, the great majority of surveys that have been carried out have been cross-sectional. Absent longitudinal data collection, it is extremely difficult to understand how health outcomes change over time and how such changes occur in the absence or presence of various lifestyle modifications.

> *Conclusion 7: Insufficient data are available on the prevalence of sleep insufficiency, economic pressures, diet, and exercise habits for the population of commercial motor vehicle drivers.*

CURRENT FATIGUE AND HEALTH AND WELLNESS MANAGEMENT PROGRAMS FOR CMV DRIVERS

Myriad types of health and wellness and fitness programs have been developed for drivers in the truck and bus industries, involving different levels of support and involvement by employers and participation by drivers. Driver health and wellness programs may be as simple as just offering health risk assessments (HRAs) and health and wellness tips and encouragement. But some employee (driver)-oriented programs include the following features: (1) conducting preemployment physicals, testing, diagnosis, and referrals for treatment; (2) employing health coaches who monitor and interact with drivers continuously while they are employed; (3) offering continuing education, hints, tips, and advice on healthy diet, weight loss, sleep and fatigue management, stress control, and lifestyle improvements; (4) conducting screening and treatment for selected sleep disorders, such as OSA (done primarily by a few large carriers); (5) offering participatory programs in health and wellness activities, such as physical fitness and exercise activities, health monitoring, and smoking cessation efforts; (6) offering motivational and incentive-based initiatives to promote employee participation in health and wellness programs, such as group games and competitions to foster weight loss and smoking cessation, providing lower-cost insurance premiums; and (7) encouraging participation in health and wellness programs by drivers' family members.

Fatigue is clearly an important concern for highway safety. Many large fleets have taken proactive measures to improve their safety figures by implementing fatigue management programs in the workplace, often incorporating fatigue management into their ongoing employee health and wellness programs. On the other hand, independent owner-operators generally must address such issues as fatigue management and health and wellness on their own (Krueger, 2012). While that important distinction exists, even fleets vary greatly in how fatigue is addressed and how health and wellness are managed. These approaches range from the minimum, which entails just complying with hours-of-service (HOS) regulations, to comprehensive programs for fatigue management and health and wellness that may include screening, training, treatment, and program evaluation.

One program, described to the panel by safety experts from Schneider National, showed that screening for OSA resulting in required use of PAP devices, if necessary, yielded positive benefits such as a substantial reduction in health care costs, a 30 percent reduction in accidents, and increased driver retention. Schneider National now prescreens all drivers for OSA, and its prescreening assessment was found to be 82 percent effective. (Some PAP devices document use, and this information can be used to terminate noncompliant drivers with OSA.)

Health and Wellness Programs of Large Carriers

While the panel did not undertake a systematic review of fatigue management or health and wellness programs instituted by large carriers, it was briefed on the experiences of three large carriers at its third meeting in September 2014. While the panel knows anecdotally that some small carriers have instituted health and wellness management programs, these appear to be few in number. This is to an extent understandable given the fixed costs of implementing such programs.

The panel heard about the employee health and wellness programs of Schneider National, J.B. Hunt, and Con-Way Motor Freight. Table 8-1 summarizes selected information from these presentations. These three carriers have integrated driver fatigue management into their health and wellness programs. Returns on investment in successful health and wellness programs include such benefits as improved employee morale and driver retention rates; enhanced corporate culture and reputation; decreased clinic visits and costs; improved insurance premium rates; and, importantly, decreased accidents and injuries, and therefore also improved safety records (Krueger, 2012).

Current Fatigue and Health and Wellness Programs in the Bus Industry

Hundreds of bus companies operate between 2 and 10 buses. Very little is known about the fatigue or health and wellness programs of these carriers. One example of a fatigue management program of a large interstate bus company was brought to the panel's attention by the chair of the Bus Industry Safety Council for the American Bus Association, Michael

TABLE 8-1 Health and Wellness Programs of Three Large Carriers as of September 2014

Component	Schneider National	J.B. Hunt	Con-Way Motor Freight
Pre- and Postjob Screening	Prework physical screen	Electronic physicals and customized tests	
Education	• Drivers and fleet managers—science of sleep/circadian rhythms (better manage fatigue) • Customers—pickup/delivery times, transit times, expectations (improve scheduling times) • Medical examiners—nature of work, sleep apnea (better detect sleep disorders) • Electronic log devices and approach to hours of service		
Training	Injury prevention training	Job safety analysis	Safety coaching, injury prevention coaching
Health and Wellness	• On-site U.S. Department of Transportation (DOT) medical examinations • Awareness programs • Availability of on-site physical therapists	Better Health for Life Program, J.B. Hunt Health Moves Program	On-site safety and health coaching
Sleep Apnea	Screening and treatment program	Sleep apnea clinical trial	Sleep study

McDonal. In his presentation to the panel, he stated that Greyhound, a large bus company, has a fatigue management initiative based on input received from focus groups comprising drivers, operations managers, and safety directors throughout the industry. On the basis of input from focus groups, Greyhound developed countermeasures to address driver stress and fatigue. Data were collected on the effectiveness of these measures from a small set of drivers during actual night runs. Based on these data, extra board operations were modified.[13]

Fatigue Outreach and Health and Wellness Initiatives of FMSCA

Since the mid-1990s, FMCSA has encouraged carriers to engage in employer-sponsored driver fatigue management programs. To this end, the agency has conducted extensive education and outreach; held forums on the topics; and carried out extensive research on driver fatigue and performance in conjunction with cooperative, enlightened truck and bus carriers. In terms of continuously addressing driver fatigue management, FMCSA developed and sponsored (with industry) the Mastering Alertness and Managing Driver Fatigue train-the-trainer 4-hour course, which was offered to more than 4,500 trucking safety and risk managers from 1996 until 2006. FMCSA also arranged to develop (with industry) an employee health, wellness, and fitness program entitled "Gettin'-In-Gear," a train-the-trainer program offered to more than 2,500 trucking safety and risk managers from 2001 to 2006.

The North American Fatigue Management Program (NAFMP), sponsored by FMCSA, Transport Canada, and others, has been available on its own dedicated website since summer 2013. This extensive set of more than 800 slides is offered in 10 training modules aimed at employers, managers, drivers, shippers, and receivers. The 10 modules are (1) introduction and overview, (2) safety culture and management practices, (3) driver education, (4) family education, (5) train the trainer, (6) truck driver safety and compliance: the role of shippers and receivers, (7) sleep disorders management for motor carriers, (8) driver sleep disorders management, (9) driver scheduling and tools, and (10) fatigue monitoring and management technologies.

NAFMP serves as an information source for the entire ground shipping industry. It offers numerous health and wellness lessons; fitness principles; and coverage of additional selected topics, especially those related to sleep hygiene, including sleep disorders such as OSA. For example, the program covers the effects of being continuously jostled, consuming a bad diet, not exercising, poor weight management, high blood pressure, using various medications, using alcohol, using tobacco, disruption of circadian rhythms, and resultant conditions such as diabetes.

As noted by Terri Hallquist of FMCSA in her presentation at the panel's third meeting, there is currently no effort under way to evaluate

---

[13] For details on Greyhound's Alertness Management Program, see http://onlinepubs.trb.org/onlinepubs/conferences/2011/FatigueInTransit/Presentations/Al%20Smith.pdf [March 2016].

who visits the NAFMP website or whether those who do visit are engaging in any behavior modification, such as sleeping longer.[14] This lack of evaluation to date is likely the result of the program's recent implementation. As discussed in Chapter 11, some evaluation of the program is feasible using web analytics. Given the attention paid to fatigue and health and wellness management by the large carriers, one key concern is whether NAFMP is adequate to educate and motivate behavior change among drivers for many midsize carriers that lack such programs or for small carriers with only a handful of trucks and drivers, or among the hundreds of thousands of independent owner-operator drivers nationwide.

> *Conclusion 8: Insufficient information exists on (1) how the variety of fatigue management and health and wellness management programs available have been designed, (2) whether drivers/employers actually adhere to these programs, and (3) whether these programs are effective in achieving their goals.*

---

[14] Note that FMCSA recently also instituted an online resource to help CMV drivers understand how a medication might impact their ability to drive safety. See http://www.fmcsa.dot.gov/medical/driver-medical-requirements/medication-issues [March 2016].

9

Technological Countermeasures for and Corporate Management of Fatigue

To repeat a key point made in Chapter 3, there is no known biological substitute for sleep; the body cannot compensate for lost sleep through any other natural mechanism. Research cited in Chapter 3 indicates that many commercial motor vehicle (CMV) drivers may be obtaining less than 6 hours of sleep per 24-hour day, which is generally viewed as being insufficient to maintain adequate levels of alertness on the job. Accordingly, as discussed in Chapter 7, sleep insufficiency increases the risk of being involved in a crash. CMV drivers also must be cognizant of circadian, time-of-day influences on their levels of alertness and the associated increased crash risk. Even when a person has obtained sufficient sleep, he or she will experience two circadian-driven lulls that adversely affect alertness, bring about feelings of fatigue, and lead to degraded performance (e.g., slowed reaction times) (see Chapter 3). To address these concerns, both technological approaches and fatigue management protocols have been advanced that are designed to detect and manage fatigue among CMV drivers.

Conclusion 9: Acute and chronic sleep insufficiency produces fatigue in drivers, as do lengthy and irregular work schedules.

Conclusion 10: There is no biological substitute for sufficient sleep.

This chapter first reviews technological approaches for detecting and managing operator fatigue. It then describes infrastructure- and vehicle-based systems designed to mitigate the effects of fatigue by alerting drivers that they are in a drowsy state. Next, the chapter briefly examines fatigue management programs. The final section addresses the importance of a safety culture.

TECHNOLOGICAL APPROACHES FOR DETECTING AND MANAGING OPERATOR FATIGUE

Fatigue detection and management approaches are of three main types: (1) online operator fatigue detection technologies, (2) fitness-for-duty indicators, and (3) biomathematical scheduling models. In general, the effectiveness of all three of these approaches in detecting or predicting operator fatigue remains unclear. Driver drowsiness detection devices and biomathematical scheduling tools are in varying stages of development (see, e.g., Mallis and James, 2012), and many of them have not been tested in third-party, randomized controlled studies. In addition, the implementation of these approaches in the field poses a variety of challenges.

Fatigue detection technologies and predictive algorithms must be shown to meet a number of criteria prior to use. First, they need to be operationally feasible, working under a variety of environmental conditions such as during driving at nighttime, on different types of roadways, in traffic, and so on. Second, they need to have small errors both when indicating that a driver's alertness is low or that he or she may be fatigued while driving and when not indicating fatigue. In other words, they need to have infrequent false positives (indicating a driver is fatigued when he or she is not) and false negatives (indicating a driver is alert when he or she is not). Third, such devices must work for a relatively heterogeneous group of drivers in a variety of driving situations, and they must designed to be user-friendly. Finally, all such devices need to be shown to measure a biobehavioral marker of hypovigilance due to drowsiness. (For details on sets of acceptance criteria for such devices, see Dinges and Mallis [1998].)

With respect to false positives and false negatives, it is clear that a system with a high rate of false negatives would be a serious problem. However, a moderate rate of false positives can also be problematic. Providing a warning to a driver when he or she is not fatigued may produce annoyance, reduce trust in the system, and decrease operator compliance with future warnings. Thus it is important that research on new technologies be focused on improving the accuracy of these devices for those that both are and are not drowsy.

It should also be noted that an effective near-real-time drowsiness detection device would be extremely useful for preventing many future crashes. But so, too, would be a device that could help in establishing whether a driver is in a sufficiently alert state before he or she begins work or during a duty day. For details, see Dinges and Mallis (1998) and Mallis and James (2012).

Online Operator Fatigue Detection Technologies

Online fatigue detection technologies are intended to provide feedback to a driver on his or her alertness level and thereby to detect operator fatigue prior to its onset. Equipped with this real-time feedback, drivers can opt to take a break from driving or use some other validated fatigue countermeasure. Additionally, drivers can monitor their alertness levels across their work-rest duty cycles to determine whether they should implement schedule-related fatigue management and mitigation strategies. Such preventive strategies include scheduling sleep and nap periods prior to the start of a duty cycle; taking rest breaks during the duty day; maximizing sleep time on nonduty days, preferably in alignment with biological night; and if possible, negotiating with employers for more fixed, regular work duty scheduling. These preventive strategies can help reduce operational errors and crashes associated with insufficient alertness attributable to neurobehavioral deficits resulting from sleep loss, circadian misalignment, or simply boredom associated with sustained work.

Online real-time operator fatigue detection technologies include a variety of methods for directly measuring factors associated with driver fatigue, such as electroencephalographic (EEG) frequency band activity; heart rate variability; and ocular variables, including saccades, slow eye movements, blink rate, or eyelid closure (see Abe et al., 2011; Chua et al., 2012; Dinges et al., 1998). One ocular measure that has consistently shown promise is PERCLOS (percentage of eyelid closure), the proportion of a time interval during which the eyes were 80 to 100 percent closed (not including blinks) (Wierwille and Ellsworth, 1994; Wierwille et al., 1994). Such measurements have been experimentally validated as being sensitive to fatigue and circadian changes (Abe et al., 2011; Chua et al., 2012; Dinges et al., 1998, 2002; Mallis and Dinges, 2005; Mallis et al., 2007; Ong et al., 2013). A PERCLOS measurement system mounted in a vehicle cab entails video monitoring of the individual, and can effectively provide the driver with an alert that he or she is becoming fatigued and suggest that it is time to take a break from driving (Mallis et al., 1998).

Early research and on-the-road testing have shown, however, that implementing PERCLOS monitoring systems for truck drivers is not always operationally feasible. Early generations of PERCLOS devices had technological limitations, as they did not accurately track slow eyelid closure throughout all environmental conditions. Some of these inaccuracies were attributable either to interference due to sunlight reflections or to difficulty in measuring pupil opening because of low light conditions during nighttime driving. Additionally, PERCLOS systems were not able to distinguish between eyes being out of field of view of the PERCLOS camera because of drivers constantly turning their head, mainly to check their side-view mirror, and drivers closing their eyelids completely. PERCLOS monitoring devices continue to be improved. Some include information on other variables to address the above challenges. For example, efforts are now under way, using optical computer recognition, to develop the capability for continuous tracking of percentage of eyelid closure in real time (Dinges et al., 2005a, 2007).

Fitness-for-Duty Testing Devices

Fitness-for duty testing devices entail probed evaluation or temporally discrete sampling of neurobehavioral performance or aspects of physiological changes (primarily ocular and pupillary measures). They represent an attempt to assess fatigue risk, especially in safety-sensitive, around-the-clock operations in which the risk of errors can be elevated because of fatigue-inducing work schedules. Unlike online fatigue detection approaches, fitness-for-duty tests produce a snapshot in time of a worker's level of alertness prior to a duty period that can predict performance capability later in the day during the duty period. The prediction is based on a comparison of baseline data specific to the individual, consisting of measurements taken when the individual was well rested or without a sleep debt.

Earlier generations of fitness-for-duty devices were relatively large, were not portable, and required relatively long sampling periods. Some were as large as the Truck Operator Proficiency System, a driving simulator installed at a truck terminal that measured psychomotor and divided-attention performance. The device was used in the 1990s in New Zealand and in Arizona. No formal predictive validity was established for the device; however, the test had high face validity, so it was believed that failing it meant the driver could not drive safety (Charlton and Ashton, 1998; Hartley et al., 2000).

Other, smaller devices used tabletop pursuit rotor tracking performance tasks, exemplified by the Factor 1000 system, which employed a critical tracking task requiring eye-hand coordination. These tabletop computerized test batteries included as many as 8 to 10 tests, including visual perception, reaction time, concentration, cognitive processing, and even personality assessment. (An example is ART 90, a fitness-for-duty test system used in several places in Europe; see Charlton and Ashton [1998]).

Most such fitness-for-duty systems were performance-based and entailed substantial learning curves because they were affected by aptitude and experience levels and did not necessarily prove to be sensitive to fatigue. Some became more useful in detecting drug effects on performance at work. Over the years, technological advances have resulted in the development of fitness-for-duty testing devices that are portable, take less time to administer, and use measures that have been demonstrated to be sensitive to fatigue. Fitness-for-duty tests that can be administered on an electronic tablet or smartphone hold particular promise for being able to predict performance capability over a brief sampling timeframe (e.g., estimating whether it is safe for a driver to extend a duty period). One caution is that the tests may not be feasible in all operational environments.

The 10-minute psychomotor vigilance test (PVT) is an example of a probed-performance fitness-for-duty test. It measures the ability of the brain to sustain attention and repeatedly respond quickly to a simple visual (or auditory) stimulus. The stimuli occur at a predetermined interstimulus interval (ISI) range, and measurements are highly precise (Basner and Dinges, 2011). There is no lengthy learning curve, and the test is independent of aptitude, experience, and skill. Reaction time, response speed, and lapses in attention (longer response times or no response) are the primary outcome measures (Basner and Dinges, 2011; Dinges and Powell, 1985). The PVT is an accepted standard for measurement in sleep and circadian research because of its demonstrated sensitivity to acute total sleep deprivation, chronic sleep restriction, and circadian rhythmicity. Research protocols conducted in operational settings also have demonstrated its validity for identifying operator fatigue (Gander et al., 2008; Pack et al., 2006; Rosekind et al., 1994)

The 10-minute duration of the PVT is often considered impractical in most operational environments because it requires the individual to disengage completely from his or her operational duties (e.g., driving) to take the test. Consequently, two shorter-duration versions of the PVT have been developed: the adaptive PVT-A (about 3 minutes, depending on the individual's performance [Basner and Dinges, 2012]) and the brief PVT-B (about 5 minutes). Both have modified algorithms for performance evaluation and have been extensively validated for their sensitivity to fatigue (Basner et al., 2011). Future studies are needed to show the feasibility and usefulness of employing PVT-like fitness-for-duty tests in safety-sensitive operational environments.

Biomathematical Models

Biomathematical models of cognitive performance and alertness are used to calculate an estimate of alertness based on sleep-wake schedules and the timing and placement of duty shifts (Dawson et al., 2011; Mallis et al., 2004). Most are based on the two-process model, as originally described by Borbély (1982). The algorithms model the known effects of the interaction between homeostatic sleep drive and the circadian system on alertness and cognitive performance. The interaction of these two physiological processes is dynamic and complex (see Chapter 3). In oversimplified terms, the sleep process is approximated to be an exponential saturation function during sleep itself, with linear degradation in performance with increasing hours of wakefulness (i.e., being awake 18 or more hours since the last sleeping period). The circadian process is of sinusoidal form and shows variations in performance on an approximately 24-hour cycle. The observed changes in performance are closely correlated with endogenous rhythms of core body temperature. These algorithms are operationally useful only when they are incorporated into a scheduling tool.

Each biomathematical scheduling tool uses specific algorithms that differ in various ways, including input variables, output measures, goals, and capabilities. It is also important to consider the operational environment for the model's intended use. For example, some models use only data collected in tightly controlled laboratory studies, while others are adjusted based on measures collected in operational environments, including military, trucking, aviation, and railway operations. The U.S. Department of Defense, U.S. Department of Transportation, and National Aeronautics and Space Administration sponsored a modeling workshop in 2002. All seven of the models evaluated at the workshop showed promise in the prediction of alertness and performance, but each required further scientific validation in tightly controlled laboratory settings and improvements in reliability, sensitivity, and specificity before being transitioned to operational settings (Mallis et al., 2004). Once such a model has been deployed in an operational setting, data collected can be used to further refine the algorithm, making it more specific to its intended use.

Biomathematical models can be used to increase safety and minimize fatigue risk by enabling comparison of different work-shift or work-rest scheduling scenarios. The goal is to maximize safety by evaluating schedule parameters scientifically rather than in accordance with time-on-task theories (see Chapter 3). Schedule evaluations include (1) predicting times when performance is optimal, (2) identifying time frames in which recovery sleep will be most restorative, and (3) determining the impact of proposed work-rest schedules on overall neurobehavioral functioning (Mallis et al., 2004).

Government entities are showing increased interest in the use of biomathematical models in the development of regulations and comprehensive fatigue risk management programs. In 2014, the Federal Aviation

Administration (FAA) made changes to the Federal Aviation Regulations (FARs) that allow carriers to submit a fatigue risk management system (FRMS) application in requesting an exemption from prescriptive regulations (FRMSs are discussed further below). Currently, the FAA uses the Sleep, Activity and Fatigue Task Effectiveness (SAFTE) model (Hursh et al., 2004) as part of its process for evaluating FRMS applications. The SAFTE model was originally developed for military and industrial settings and is the foundation of the Fatigue Avoidance Scheduling Tool (FAST) (McCormick et al., 2013). And the U.S. Navy supported development of the OWL (Optimized Watchbill and Logistics) tool based on a published biomathematical model (McCauley et al., 2009, 2013) to generate duty schedules for sailors that minimize fatigue risk.

Currently, at least one biomathematical model has shown evidence of being able to predict cognitive deficits associated with chronic sleep loss (McCauley et al., 2009, 2013). Although improvements in biomathematical model development continue, however, limitations still exist. Scheduling tools that incorporate biomathematical models of fatigue and alertness are only one component of a comprehensive fatigue risk management program, and such models are suitable only for use as a tool to help inform final decisions about a schedule's level of safety. Moreover, only one published model incorporates individual differences, providing more accurate estimates of fatigue (Van Dongen et al., 2007). Refinements in biomathematical models to account for individual differences would facilitate the use of individual countermeasures (e.g., recovery sleep, naps, or caffeine).

Conclusion 11: Operator fatigue has been singled out for its negative safety implications for all workers, including commercial motor vehicle drivers. Such concerns have motivated a variety of applied research projects on detecting, preventing, and managing fatigue.

Conclusion 12: Despite almost three decades of research on the topic, technological innovations for detecting driver fatigue in near real time and operational strategies for their use are still in the early phases of understanding and application.

Conclusion 13: Biomathematical models can be useful for the development of general work-rest schedules. However, existing models do not account for individual variation, so care must be taken in applying them to address likely impacts of irregular work schedules.

SYSTEMS DESIGNED TO MITIGATE THE EFFECTS OF FATIGUE

Both road infrastructure-based and vehicle-based systems have been developed to mitigate the effects of fatigue by awakening drivers or warning them that they are in a compromised state and at risk of an accident.

Rumble Strips

Roadway rumble strips are an infrastructure-based system that does not specifically target driver fatigue, but may be effective in helping to prevent a range of crashes related to taking one's eyes off the road as a result of either driver fatigue or distracted driving. Rumble strips may alert a sleepy driver as well as redirect the attention of an alert but distracted driver.

Studies of the safety effect of rumble strips consistently have shown a reduction in run-off-the-road crashes, head-on collisions, and sideswipes, crash types thought to be related to fatigue and distraction. One study on rumble strips in Maine directly measured the impact on crashes identified as fatigue-related in crash reports, estimated as a reduction of 58 percent in fatigue-related run-off-the-road crashes on rural freeways (Garder and Davies, 2006).

Most studies of rumble strips have focused on crashes in specific states. Most have been before-and-after studies with comparison groups to control for exposure and various other potentially confounding factors, such as shoulder width and roadway curvature. Resulting estimates of effectiveness cover a broad range, from about 12 percent on freeways with speed limits under 65 mph to more than 50 percent for run-off-the-road fatal crashes. While the range of effectiveness estimates may be broad, it is important to note that all the estimates are positive and show a significant reduction in lane- and road-departure crashes (Datta et al., 2015; Khan et al., 2015; Patel et al., 2007; Sayed et al., 2010).

One study of Connecticut roads specifically addressed the possibility of fatigued drivers using the rumble strips to alert them repeatedly to their own drowsiness and redirect their attention, thus encouraging them to keep driving until they drifted off the road along a stretch with no rumble strips. The study uncovered some evidence for this phenomenon, finding that some stretches of road with no rumble strips experienced an increase in run-off-the-road crashes relative to stretches of road along the same route with rumble strips installed (Smith and Ivan, 2005). A simulator study of fatigued drivers showed how this might occur. Thirty-five night shift workers were tested immediately after finishing work with a simulated driving task along a two-lane road. Objective measures of sleepiness included EEG reading and eye-closure duration. All drivers showed an increase in sleepiness immediately prior to hitting the (simulated) rumble strips. The rumble strips had the effect of alerting the drivers, as indicated by the objective measures, but the drivers returned to their baseline fatigue within about 5 minutes (Anund et al., 2008).

Thus, one drawback of rumble strips is that they may provide drivers with a false sense of security. Even if drivers feel drowsy, they may not pull over and stop to take a rest break, thinking that rumble strips will continue to alert them, and so they become increasingly at risk the longer they continue to drive. As a practical matter, installing roadway rumble strips requires a costly capital investment in highway infrastructure and sustained maintenance of the system.

> *Conclusion 14: Roadway rumble strips serve to help prevent driver fatigue-related accidents. At the same time, there is the danger that drivers will treat rumble strips as if they provide repeated emergency alarm protection from falling asleep, and therefore will postpone taking other valid fatigue countermeasures, such as stopping for a rest.*

Rest Areas

When CMV drivers reach the limit of a duty period in accordance with hours-of-service (HOS) regulations, they are required to stop driving for a 10-hour stretch and spend at least 8 hours in the sleeper berth. Designated public rest areas are constructed along Interstate highways to provide both car and truck/bus drivers with a safe location to pull over and rest. Many such public rest areas have parking lots specifically allocated for trucks and buses. On many routes, moreover, numerous commercially operated truck refueling rest stops are available to accommodate many of the needs of long-haul truck drivers, including restaurant food, convenience store purchases, lounges, bath and shower rooms, and so on. These rest stops also provide spaces for extended truck parking to permit drivers to sleep in their truck-mounted sleeper berth.

Rest stops play a crucial role in contributing to highway safety by allowing fatigued drivers to pull off the road. McArthur and colleagues (2013) investigated the safety implications of rest areas by analyzing the relationship between the proximity of a road segment to a rest area and the frequency of fatigue-related crashes occurring on that road segment. Between 2006 and 2010, the authors collected crash data on all road segments that fell within a 20-mile radius of rest areas in the state of Michigan. Their spatial analysis provided evidence for the safety benefit of rest areas in reducing the frequency of fatigue-related crashes. Another finding of the study was diminishing returns from rest areas appearing as the 20-mile radius was expanded, pointing to the importance of spacing of rest areas. State-specific studies (Banerjee et al., 2009; Carson et al., 2011; SRF Consulting Group, Inc., 2007; Taylor et al., 1999) likewise have shown a positive relationship between spacing of rest areas and fatigue-related crashes. All these studies found that rest area spacing of more than 30 miles led to increased crash risk.

Two of the studies also found that overcrowded rest areas or those with insufficient parking spaces for trucks were positively associated with increased crash rates (Banerjee et al., 2009; Carson et al., 2011). In 1996, the Federal Highway Administration (FHWA) commissioned a study to evaluate truck driver rest and parking needs along the National Interstate System. The study identified a shortfall of 28,400 truck parking spaces in public rest areas nationwide and predicted the shortage would grow to 39,000 in 10 years (U.S. Federal Highway Administration, 1996).

One way of addressing this problem is technology that can inform drivers where to anticipate a parking place where they can sleep. A large number of public/municipal parking garages (e.g., at airports or shopping centers) currently identify electronically at the front entrance not only how many parking spaces are available but also on which floor and even the specific spaces that are available. It would appear that related technological innovations could be applied to on-the-roadway parking facilities as well. The American Transportation Research Institute (ATRI), in collaboration with the Minnesota Department of Transportation, is developing a system that can identify available trucking spaces and communicate that information to drivers.[1]

It should be pointed out as well that constructing, maintaining, and policing the security of public parking for CMV driver rest areas has significant cost implications for federal, state, and municipal governments. Thus the trucking industry may need to take a leadership role in generating more parking at commercially operated truck refueling rest stops in geographic areas where it is needed the most.

> *Conclusion 15: Repeated surveys by trucking industry and other research organizations have revealed insufficient numbers of publically available rest areas where commercial motor vehicle drivers can safely take a lengthy rest. This issue has the potential to impact fatigue-related crash risks and needs to be addressed by the safety community.*

Vehicle-to-Infrastructure Systems

Vehicle-to-infrastructure (V2I) technology allows for wireless communication between passenger vehicles and traffic and highway infra-

---

[1] Truck Parking Availability Study: Demonstration Project. Available: http://atri-online.org/wp-content/uploads/2012/04/truckparkingonepager.pdf [March 2016].

structure to prevent collisions and manage traffic. Examples of V2I systems include monitors on bridges that communicate ice accumulation to approaching vehicles, traffic signals that warn vehicles of stopped traffic, and sensors warning of nearby emergency vehicles or work zones. A National Highway Traffic Safety Administration (NHTSA) analysis found that such technologies could potentially prevent 26 percent of all vehicle crashes (Najm et al., 2010). Even though they are not designed to address driver fatigue, such technologies that warn of traffic slowness/stoppages could also reduce fatigue-related crashes, thereby making them an effective fatigue countermeasure in certain circumstances. Whether V21 systems that warn of traffic slowness/stoppages reduce fatigue-related crashes is a topic for further research.

Vehicle-based Systems

In addition to the systems installed on the vehicle to detect drowsiness/fatigue in drivers discussed earlier, vehicle-based systems for reducing fatigue-related crash risk include forward collision warning, automatic emergency braking, lane departure warning systems (LDWS), blind-spot object detection, and adaptive cruise control. Other systems monitor the driver's use of controls, such as steering and braking, to detect degraded performance. Changes in the pattern of steering wheel adjustments, for example, have been used to detect degraded driver vigilance.

LDWS were developed for use on heavy commercial trucks as a type of driver alertness monitoring system. These systems are designed to warn drivers when they drift from their driving lane unintentionally, perhaps as a result of driver fatigue or distraction.

LDWS include several types of sensors installed on vehicles to monitor lane-tracking/lane-keeping performance, and then provide warnings when drivers deviate noticeably from the center of the lane over time (i.e., the past several minutes). Video sensors—usually camera-like devices mounted behind the windshield and aimed at the roadway in front of the vehicle or integrated beside the rear view mirrors—detect visible roadway lane and edge painted markings. These sensor data are fed into onboard computerized image recognition software to track a driver's lane travel performance. Infrared sensors (either behind the windshield or under the vehicle) may be part of the system as well. The video sensors also may be accompanied by laser sensors mounted on the front of the vehicle.

Lane drift can be the result of anything from drowsiness and distraction to adverse weather conditions (e.g., snow, rain) that obscure the roadway paint markings. The LDWS continually monitor a vehicle's position and detects when the vehicle begins to drift toward an unintended lane change (e.g., perhaps approaching the roadway edge near the shoulder). Drivers inform the LDWS of plans to make intentional lane changes, such as to pass another vehicle, by first activating the turn signal device. Upon detection of lane drifting, the LDWS may present the driver with a visual display of his or her lane-tracking performance and/or present an audio warning. In the case of the AutoVue® LDWS (developed by Iteris Corporation), when a driver drifts out of a lane, the system emits a distinctive "rumble strip" noise from right or left door speakers. The LDWS may present other audible warnings to alert the driver to make a course correction to stay within the lane. False alarms are minimized by disabling the warnings when the vehicle's speed is low.

A few versions of LDWS, particularly early prototype systems, were subjected to independent on-the-road testing in heavy trucks (see, e.g., Dinges et al., 2005a). Although LDWS were not completely validated as a driver fatigue predictor in such road testing (it would be unethical to induce driver fatigue for such testing), such a system was shown to "sharpen lane position awareness." Moreover, LDWS gained acceptance by drivers as alertness monitors, and it was noted that in some cases, they even helped to improve driving skills.

The Iteris AutoVue LDWS was fielded on Mercedes Actros commercial trucks in Europe as early as 2000. In 2002, Freightliner Trucks' North American vehicles made the Iteris system available as an "after market" option. Before selling the AutoVue system to Bendix CVS in 2011, Iteris reported that the system was in use in thousands of trucks, sold as an original equipment manufacturer option on new class 8 trucks. The system is widely available in newer automobile models as part of special option order electronic monitoring packages. It is clear that such LDWS technology could eventually allow truck fleet owners to analyze near-real-time safety information transmitted wirelessly from their vehicles using existing fleet communication systems.

Partially Autonomous Vehicles

In the relatively near future, partially autonomous vehicles may be used widely. Such vehicles are unlikely to fully replace the driver in the near term because driving remains a complex task. Nonetheless, partially autonomous vehicles have the potential to reduce the likelihood of fatigued driving by performing a majority of driving tasks (without direct involvement of the driver), thus reducing the driver's cognitive workload and attention-demanding tasks.

On the other hand, partially autonomous vehicles could have a fatigue-related negative impact on safety, especially in driving situations in which the hand-off of control to the human driver went beyond the design parameters of the system. Basically, as soon as the driving task became complicated, such as when the vehicle was entering a dense traffic area or something unexpected happened, the driver would need to take back control from the vehicle. How this would be done safely is both a design and operational challenge. Currently, autonomous driving systems do the easy part—driving vehicles down the road where the main challenge is lane keeping. But should a challenge arise that required the driver to assume control of the vehicle, that driver could have been lulled into disengagement by the system and could, for example, be using a smartphone or a laptop computer.

The success or failure of autonomous driving systems relies on effective human-system integration in design and practice. For such automation to be successful, the human user must be aware of the automation and react to it appropriately (see Shaikh and Krishnan, 2012). The type of warning that is most effective in attracting a driver's attention and at times alerting the driver to the need to retake control of the vehicle needs to be determined.

Given that most of these systems may have sensors that can relay information from the vehicle to dispatchers or fleet managers, carriers might like to have the ability to force a truck or bus driver to pull over before a situation resulted in a crash. In any case, since the decision-making abilities of a fatigued driver are compromised, relying on the driver's decision to pull over might not be sufficient in many cases. This is one of the many reasons why fatigue detection and mitigation technologies and autonomous driving systems need to be carefully designed, then thoroughly tested, evaluated, and validated. Moreover, given the various incentives that affect a driver's decision making concerning pulling over to take a rest when tired, there is likely an important benefit to testing such systems in less controlled settings, such as in naturalistic driving studies.

Additional Research

NHTSA is exploring new on-board technologies for combatting drowsy driving. DrIIVE (Driver Monitoring of Inattention and Impairment using Vehicle Equipment) is an NHTSA project currently under way. Drowsiness often is evidenced by short episodes of degraded performance. Thus, the goal of the project is to use vehicle-based driver behavior data to predict impairment due to alcohol, drowsiness, and distraction (Brown et al., 2014). The research entails examining steering and pedal inputs and lane position and comparing this information with "signatures" of normal driving, when a person is awake and alert. Another goal of the project is to demonstrate the potential to detect these states without the use of cameras to monitor drivers' faces.[2]

In addition to the crash avoidance technologies available to truck drivers and fleets and work being pursued by NHTSA, car manufacturing companies have developed systems to warn drivers of inattentiveness or drowsiness. Mercedes-Benz's Attention Assist system helps drivers recognize when they are drowsy or inattentive and advises them to take a break. When drivers are alert, they constantly, and subconsciously, monitor the position of their car and make continual small steering adjustments to keep the vehicle on a safe path. When drivers are fatigued, they experience periods of inattentiveness during which there is little steering input, followed by sudden and exaggerated corrections when the driver regains attention. Attention Assist uses a sensitive steering angle sensor to monitor the way in which the driver is controlling the car. At speeds between 80 and 180 km/h (55 to 110 mph), the system identifies a steering pattern that is characteristic of drowsy driving and combines this with other information, such as time of day and duration of journey. If a sequence of such events is identified, the system warns the driver to take a break by showing a coffee cup signal on the dashboard and emitting an audible tone. The driver may acknowledge the warning to make it disappear from the display. If the driver does not take a break and his or her driving style continues to indicate drowsiness or inattentiveness, the warning is repeated after 15 minutes.

Bosch is designing a system to evaluate driver microsleep, determine the level of drowsiness, and warn the driver if necessary. The system analyzes steering behavior to identify when the driver does not steer and then makes an abrupt steering correction. This system also makes use of the speed and duration of travel.

Finally, it is worth mentioning, as is the case for rumble strips, that safety issues can arise if crash avoidance technologies are considered countermeasures for fatigue because they provide protection against crashes, rather than addressing fatigue. These devices essentially help protect drivers from some of the consequences of fatigue, but they should not be used as a justification to continue to drive.

---

[2] Previous contributions by NHTSA include detection of impairment from alcohol using vehicle measures (DOT HS 811 358), visual measures for detecting driver distraction (DOT HS 811 547A), and advanced countermeasures for multiple impairment (DOT HS 811 886). The agency is in the process of (1) developing and evaluating a system of algorithms for identifying signatures of alcohol-impaired, drowsy, and distracted driving; (2) assessing potential countermeasures for drowsy driving-associated lane departures; and (3) evaluating an initial proof of concept for the use of this system in the development of safety technologies such as driver feedback displays for drowsiness.

Conclusion 16: Additional research is needed on the effectiveness of all devices that may address reduced vigilance due to fatigue, including forward collision warning, automatic emergency braking, lane departure warning systems, blind-spot object detection, adaptive cruise control, and any other in-vehicle driver drowsiness/fatigue detection systems. This research needs to encompass not only the devices' effectiveness but also the results of actual deployment, the impact of driver acceptance, and any negative consequences of using such devices inappropriately as countermeasures for fatigue.

FATIGUE MANAGEMENT PROGRAMS

Fatigue management programs define policies and procedures for managing and mitigating fatigue in safety-sensitive environments (Lerman et al., 2012). They often are implemented within health and wellness programs or safety management systems. Each freight- or passenger-moving operation is unique and presents its own fatigue challenges. A key feature of fatigue management programs is that they consider both physiological (see Chapter 3) and operational factors. Therefore, these programs most commonly are tailored to the operational needs and constraints of particular work settings (or companies) (Dinges and Mallis, 1998; Mallis and James, 2012).

Fatigue management programs can be classified into two broad categories: (1) fatigue risk management plans (FRMPs) and (2) fatigue risk management systems (FRMSs). FRMPs establish policies on managing and mitigating fatigue during operations. They typically include a requirement for fatigue awareness training for employees (e.g., drivers, fleet managers, dispatchers), as well as processes for reporting instances of fatigued driving. FRMSs take FRMPs one step further as they aim to manage operator fatigue at a more granular level. They include a continuous feedback loop that provides a means for continuous measurement and monitoring of an individual worker's schedules using subjective and objective data collection (see Gander [2015] for additional details, and Fourie et al. [2010a, 2010b] for discussions of the effectiveness of FRMPs and FRMSs in the trucking industry.)

Finally, the Federal Motor Carrier Safety Administration (FMCSA) and Transport Canada have worked over the past decade to develop the North America Fatigue Management Program (NAFMP), discussed in detail in Chapter 8. The purpose of this online program is to present effective ways to manage and mitigate fatigue in trucking operations. The information provided encompasses prescriptive fatigue-related regulations (i.e., HOS rules), fatigue awareness, the nature of sleep and sleep disorders, work-rest schedule development, and known fatigue countermeasures.

> *Conclusion 17: Fatigue risk management plans and fatigue risk management systems used in aviation, the rail industry, and the pipeline industry need to be studied further since they may provide models that can be applied to commercial motor vehicle driving.*

IMPORTANCE OF SAFETY CULTURE

This chapter has described various measures and technologies for dealing with CMV driver fatigue, all of which are aimed at the driver. Drivers play a major role in safe CMV operations, and if they are more aware of their degree of fatigue and how best to counter it, the risk of crashes should be reduced. Barriers to entering the CMV driver profession are somewhat low, since all one must do is obtain a commercial driver's license (CDL) (see Chapter 2 for description of different licensing requirements) and pass the U.S. Department of Transportation's (DOT) physical exam.[3] Trucking companies do prefer a clean driving record when hiring drivers, and many of them hire from CDL training schools. Safety training is a significant part of the curriculum of these schools (e.g., how to drive on ice or snow or in the rain). However, the curriculum does not always include sufficient coverage of driver fatigue awareness. Therefore, one cannot be certain that a qualified truck driver is fully aware of the risks of driver fatigue and its consequences. Big trucking companies usually have orientation and on-the-job training programs that educate drivers on fatigue and how to manage it. Such companies are aiming to ensure that their employees are safe drivers, thereby maintaining their safety records and controlling one of the most obvious costs—from accidents.

Focusing on corporate safety records has its benefits, as safety and economic gains are linked—unsafe drivers are bad for business. Both economic gains and a company's approach to safety are determined by the company's organizational culture—values and norms held and shared by workers on the aims of the company and how the work should be done. FMCSA requested that the Transportation Research Board's Commercial Truck and Bus Safety Synthesis Program (CTBSSP) broaden the understanding of "safety culture" and synthesize best practices and guidelines on the development of such a culture among motor carriers. *CTBSSP Synthesis 14* (Short et al., 2007) highlights the importance of treating safety as the responsibility not only of drivers but also of dispatchers and fleet managers, as they are responsible for scheduling loads and are aware of how many hours drivers have been on duty. *CTBSSP Synthesis 14* cites

---

[3] DOT Medical Exam and Commercial Motor Vehicle Certification. See https://www.fmcsa.dot.gov/medical/driver-medical-requirements/dot-medical-exam-and-commercial-motor-vehicle-certification [March 2016].

examples of trucking companies undertaking initiatives on various fronts that mirror these best practices and guidelines.[4] In essence, it is necessary to educate dispatchers, fleet managers, and safety managers about the fatigue-related challenges faced by drivers.[5] Fatigue management programs, in combination with training for drivers and trucking officials, can help drivers make optimal use of their off-duty hours.

As discussed in Chapter 2, the trucking industry is highly heterogeneous in terms of operational structure; certain populations of drivers are not formally taught a wide array of safe driving practices, and driver fatigue management may or may not be part of that training. When independent owner-operators engage in sustained contract work for larger carriers, they occasionally are expected to engage in the same training received by the company's employees. Generally, however, independent owner-operators lack ready access to the same depth of education and training available to drivers working for large carriers.[6] However, they do have ready access to the NAFMP online. In the end, however, implementation of fatigue management practices depends mainly on personal initiative by the independent driver.

The above discussion relates to the concept of safety culture, which is achieved when shared values and beliefs interact with a carrier's structures to produce behavioral norms. It is important to study how the different approaches to safety culture of various carriers relate to the decisions made by CMV drivers about whether to continue driving when they feel fatigued. A separate but related concept of "safety climate" is also worthy of study for its impact on driver behavior. Safety climate is defined as "shared perceptions of the organization's policies, procedures and practices as they relate to the value and importance of safety within the organization. In short, safety climate is the measurable aspect of safety culture" (Huang et al., 2011).

> *Conclusion 18: Further research is needed on the impact of safety culture on driver decision making with respect to countering fatigued driving and on crash frequency.*

---

[4] See http://www.truckinginfo.com/article/story/2012/10/building-a-culture-of-safety.aspx [March 2016].

[5] There is a precedent for educating motor carriers and their drivers about fatigue—the Mastering Alertness and Managing Driver Fatigue Program run by FMCSA and ATRI, the research arm of the American Trucking Associations.

[6] Some independent owner-operators who provide services to large trucking fleets that have fatigue management programs may have access to the fleet's training and education programs.

PART IV

RESEARCH DIRECTIONS

10

Research Directions for Fatigue and Highway Safety

The causes of a crash can be related to the characteristics of the driver; driver behavior/performance; the vehicle; the fleet; or the environment in which the driver and the vehicle, as well as other vehicles, were traveling. For a given crash, multiple factors can and often do play a part. Any analysis of individual causal factors, including driver fatigue, can be biased by a failure to represent the impacts of various confounding influences. Therefore, assessing the role of driver fatigue in highway safety requires a comprehensive approach that accounts for the contribution of all the important factors that can cause changes in the degree of driver fatigue and in crash frequency. To this end, one must have data on the contributions of these other causal factors.

This need in turn prompts the need for an assessment of what data on the various causal factors exist, at what level, and for what populations and how linkable these data are to other variables. Such an assessment can provide a better understanding of the existing data gaps and how they might be filled to develop a comprehensive database that can support more conclusive research on the role of driver fatigue and hours-of-service (HOS) regulations in highway safety. With such a database, research would be able to identify which factors play more or less important roles in causing driver fatigue, and which factors, including fatigue, play more or less important roles in the frequency of crashes. Along with other considerations, such as the feasibility of implementing policy changes that can alter various causal factors, such research should greatly assist the Federal Motor Carrier Safety Administration (FMCSA) and other government agencies in determining how to focus their efforts to reduce crashes due to fatigued driving and, more generally, to improve highway safety.

Specifically, one needs to collect data from commercial motor vehicle (CMV) drivers on their degree of fatigue (e.g., hours of service, work hours, work demands); other driver factors, such as medication use, drug use, and body mass index (BMI); vehicle factors (e.g., noise, distraction, quality of brakes, visibility); carrier factors (e.g., scheduling, compensation methods); and environmental factors (e.g., time of day, road type, traffic conditions). Table 10-1 represents an initial attempt to outline the various predictors that might be expected to have an association with crashes or other safety outcomes. The table includes driver fatigue as both a predictor and an outcome because it is necessary to understand not only the extent to which driver fatigue causes crashes but also, assuming that fatigue is an important causal factor for crashes, what factors cause drivers to be fatigued.

The table also includes traffic density and whether driving occurs during the nighttime as important factors in crash risk. Both are linked to increased crash risk. Moreover, since these two factors tend to be negatively correlated, their joint impact on crash risk can be difficult to anticipate. That is, the time that is worst for staying awake—nighttime—is the best for avoidance of other vehicular traffic. Therefore, any assessment of crash risk needs to include the contribution of both factors. This is a simple example arguing for a comprehensive approach to the question of fatigue as a causal factor for CMV crashes.

Several points need to be emphasized about Table 10-1. First, it is incomplete as to the potential factors that might be included. Confirmatory research has not been conducted to enable a comprehensive understanding of the causal structure underlying crash risk. The table is also incomplete regarding data sources and their availability. Finally, it is incomplete with respect to the information that one might want to know about each data source. Completing the column on predictors is a challenge, but for existing data sources, columns 4, 5, and 6 could easily be filled in. The panel believes FMCSA can use this table as a starting point for a living compendium of the factors that affect crash risk and the information available about them. FMSCA is best positioned to complete the table because it has the most thorough understanding of the various data sets and what information they do and do not include. Also, analysts using these data to draw inferences about fatigue, hours of service, and highway safety may find that the table needs to be augmented to include columns representing other important data features.

It is also important to note that determining which factors can raise or lower crash risk is important in and of itself, and is necessary to support further causally relevant research on the interrelationship among fatigue, hours of service, and crash risk. There is a chicken-and-egg problem here in that until one knows which factors to include as confounders, one cannot know whether a variable is in fact a causal factor or simply correlated with a true causal factor. The panel believes that for now, it is better to err on the side of including variables whose status regarding causal impact on crash risk is unclear.

The column labeled "Public/Private" is a reminder that much of the data collected now is not easily available to researchers (see Chapter 5). In addition, as the columns "Level of Aggregation" and "Populations for Which Available" suggest, the data often are not available at the necessary level of aggregation or for an appropriately representative population of CMV drivers. Relevant to level of aggregation, some of the predictors will need to be linked to each trip sequence (possibly every few minutes of each trip) and so at a very detailed level. Examples include degree of driver fatigue, degree of precipitation, and traffic density. However, some other variables, such as all fleet variables and many environmental factors, including number of lanes and geometry of the road, are stable over long periods of time and can be collected infrequently.

Various definitional and measurement complications result from the need to represent many of the above concepts. With respect to outcomes, what is a crash? Does one count curb strikes? For serious crashes, does one require a threshold on damages, and should one use only avoidable crashes? Should the focus be on crashes that result in fatalities? What safety-critical events (SCEs) are relevant to analysis of fatigue and highway safety? This lack of clarity in measurement is true for some of the predictors as well. For example, previous chapters have addressed the difficulty of defining and measuring driver fatigue, but how should safety culture be defined and measured? While additional work in this area is needed, definitional vagueness and measurement complexity often can be assigned lower priority since inferences are frequently robust to the precise definitions and measurements used.

Filling out the remaining columns of Table 10-1 would clarify what data are available to FMCSA and academic/industry researchers and identify existing data gaps. FMCSA then would need to determine the relative priority of each gap and the best means of closing these gaps. FMCSA would need to determine whether private sources exist, and if so, whether various techniques could be effective for making such sources public while avoiding disclosure of individual data. Further, FMCSA would need to determine the level of aggregation at which the data exist and whether they are sufficiently detailed, as well as whether the data exist for a subset or for the entire population of CMV drivers.

Finally, the data need to be linked so that rapidly changing information on drivers (e.g., whether they obtained sufficient sleep) and the

TABLE 10-1 Factors Associated with Outcomes Representing Crash Risk

Predictor Domain	Predictors/ Variable Set	Data Source	Public/ Private	Level of Aggregation	Populations for Which Available	Outcomes
Driver	• Demographics (age, race, gender) • Health conditions (body mass index, hypertension, apnea) • Medications used • Fatigue (hours driving, hours on the job) • Recent sleep history • Circadian effects • Driving experience • Safety record (number of lane deviations, unusual speed changes, unusual brake applications) • Decision making • Work demands • Exposure	• Bureau of Labor Statistics • Driver surveys				• Crash rate • Serious crash rate • Fatal crash rate • Safety-critical event rate • Fatigue
Trucks and Buses	• Type and age of truck/bus • Quality of brakes • Quality of tires • Other mechanical condition • Maintenance frequency and history • Vehicle crash history • Technology on board for distraction avoidance	• Carriers				

Driving Environment	• Weather • Degree of precipitation • Time of day • Traffic density • Road type • Degree of road lighting • Hazards • Safety features • Availability of rest stops • Impact of other drivers • Amount of sunlight	• Police accident reports • Accident reconstruction reports
Carrier	• Operation type • Fleet size • Scheduling • Logistics • Driver turnover rate • Fatigue management program • Safety culture • Safety record • Compensation (general method, level of compensation, and compensation by various tasks)	

environment (e.g., degree of precipitation) can be merged in a way that is amenable to statistical modeling of the most relevant information. (The assumption here is that carrier and vehicle data are less changeable, but there are likely exceptions to that assumption.) Linking such information would be greatly facilitated by some means of identifying unique drivers and trips that could be used across databases, since the information described here is not likely to reside in a single database.[1]

The fact that FMCSA will discover many data gaps should be no surprise. As is clear from a review of Chapter 5, none of the data sets currently available for assessment of the role of driver fatigue in highway safety provides all the information necessary to draw rigorous inferences for the full population of interest. *Crash data sets* underrepresent the degree to which fatigue contributes to crashes since they are based on after-the-fact police reports and therefore include incomplete information on the driver's recent sleep history. Further, since comprehensive exposure data are not available, crash counts cannot be translated into crash rates and therefore support risk assessments for individual factors. *Naturalistic driving data sets* typically include few crashes because crashes are rare events. Moreover, since all such research requires participants' informed consent, it collects data only on volunteers, who may be more or less subject to certain risk factors relative to the general CMV driver population. Also, the various *special-purpose data sets* comprising small surveys, driver logs, output from various devices, and data collections for particular populations (e.g., drivers for large fleets) are available only for subpopulations and for a subset of the necessary causal factors. As discussed in this chapter, new data collection efforts and new technologies for automatically collecting relevant data could provide much of the needed data not currently available.

The remainder of this chapter presents the panel's analysis and recommendations with respect to the most important directions for future research on fatigue among CMV drivers and highway safety. It describes in turn (1) survey data that could be collected on CMV drivers to help reduce some of the major data gaps on drivers that inhibit this research; (2) data available from vehicles themselves that could help close existing gaps in data on drivers, vehicles, and the environment; (3) other data sources that could be tapped for this research; (4) examples of key research questions that could be investigated with better data; and (5) methodological and statistical issues entailed in research on driver fatigue and highway safety and how these issues can be addressed.

---

[1] Longitudinal Employer-Household Dynamics (LEHD) is an example of a data set in which employer-employee information is linked and therefore is useful for tracing the employment history of a worker. See http://lehd.did.census.gov/ [June 2016].

COLLECTION OF SURVEY DATA ON CMV DRIVERS

The population of CMV drivers in the United States is large and diverse (see Chapter 2). Currently, there are approximately 3 million CMV drivers (depending on the definition used) for whom knowledge remains incomplete.[2] Given that FMCSA is charged with establishing HOS regulations, setting standards for medical certification, and communicating the dangers of driver fatigue, the agency would greatly benefit from knowing important characteristics about this population. For instance, truck drivers who drive locally on set routes are less likely than long-haul drivers to be affected by changes to HOS regulations. Given the lack of a continuing survey, however, much about the population of CMV drivers remains undocumented, even on matters of simple demographics. This includes the number of drivers engaged in various types of employment (e.g., the number of local versus long-distance drivers); the number of hours in a day and number of days in a week that drivers with various types of employment drive; how drivers are compensated; and their age, gender, and race. This information is lacking in particular for the large number of independent drivers who are not directly employed by carriers. Absent this information, it is difficult to estimate accurately how many drivers are directly affected by FMCSA's guidelines and policies and how many might benefit from various proposed changes.

Capturing many of the measurements needed for research on fatigue and on the health and welfare of CMV drivers can be viewed as somewhat intrusive. Such data collection often is dependent upon CMV drivers who volunteer this kind of personal information. It is therefore common for various subsets of the truck and bus driver population to be underrepresented in such research efforts. Having information on the total CMV driver population and some of its characteristics would support appropriate weighting of samples so that findings could be generalized to the full population of drivers (based on certain assumptions).

Commercial Driver's License Information System Database

Generally speaking, a person wishing to legally operate a vehicle with a gross vehicle weight rating (GVWR) of greater than or equal to 26,001 lb or to transport 16 or more passengers needs to have a commercial driver's license (CDL). Obtaining a CDL requires passing a written test and a road test. Information on drivers with CDLs, along with their driving

---

[2] For purposes of issuing a CDL and for drug testing, one includes only drivers of vehicles greater than 26,000 lb gross vehicle weight rating. For medical qualification, the definition is based on FMCSA regulation 390.5 and is for drivers of vehicles greater than 10,000 lb gross vehicle weight rating.

records, is collected by the individual states and aggregated in the Commercial Driver's License Information System (CDLIS). The purpose of this database is to serve as a clearinghouse and repository of information pertaining to the licensing and identification of CMV drivers. State driver licensing agencies use the CDLIS to transmit information on out-of-state convictions, transfer the driver record when a license holder moves to another state, and respond to requests for driver status and history. Prior to the development of the CDLIS, drivers could obtain multiple CDLs and could use that capability to conceal violations from law enforcement personnel or prospective employers. Each state separately administers its own portion of the CDLIS, and all of the system's files are linked together in a national relational database.

In addition to information on a driver's safety record, the CDLIS contains information on a driver's age, gender, height, and weight. As of February 2008, the CDLIS contained more than 13 million CMV driver records (see Federal Motor Carrier Safety Administration, 2008). Since it is generally believed that there are about 3 million active truck and bus drivers, the CDLIS contains a substantial number of records for individuals who are no longer active CMV drivers. In addition, the currency of many of the addresses in the database is a concern. (As noted above, part of the vagueness comes from how a CMV driver is defined. Sections 383.5 and 390.5 of the FMSCA Safety Regulations differ in their definitions.) In addition, given that access to this database is limited to FMCSA and its contractors, using it either as a sampling frame or to compute demographic summary statistics (age, gender, race) would raise privacy concerns. Thus, while the CDLIS appears to offer a source of data on CMV drivers, it has a number of limitations that constrain its use for gathering information on CMV drivers and crash risk.

Commercial Motor Vehicle Driver Surveys

Surveys of the population of CMV drivers have been attempted. Recently, the National Institute on Occupational Safety and Health (NIOSH) conducted a survey of long-haul truck drivers. The survey design involved randomly selecting limited-access highway segments and then randomly selecting truck stops within those segments. Truck drivers entering those truck stops during a 3-day interview period were recruited for the survey. Drivers also could be approached at fueling stops or weigh stations. This survey was limited to long-haul truck drivers. (See Sieber et al. [2014] for details.) Another survey was recently carried out as part of the Behavioral Risk Factor Surveillance System in Washington State (for details, see Bonauto et al. [2014]). Respondents for this survey were contacted via telephone, which raises the concern that long-haul truck drivers were underrepresented. This is hardly a comprehensive list of surveys of CMV drivers, but it is safe to say that such studies are not carried out on a regular basis, and they generally are available only for specific subgroups.

Until 2002, the Census Bureau conducted a Truck (or Vehicle) Inventory and Use Survey,[3] which collected some of the information needed to populate Table 10-1. The sampling frame was constructed from files of truck registrations identified as being active. Sampling was stratified by state and by truck characteristics. Body type and GVWR determined the following five truck strata: (1) pickups; (2) minivans, other light vans, and sport-utilities; (3) light single-unit trucks (GVWR under 26,000 lb); (4) heavy single-unit trucks (GVWR over 26,000 lb); and (5) truck-tractors. Within each stratum, a simple random sample of truck registrations was selected without replacement, producing a sample of approximately 136,000 truck registrations. Questionnaires were mailed to the addresses corresponding to these registrations, and the results were tabulated. This survey provided information relevant to the present study, including whether a vehicle had been involved in a crash, the vehicle type, the jurisdiction, the type and size of the carrier, the distance traveled, and the range of operation. Unfortunately, this Census Bureau survey was discontinued after 2002.[4]

The Need for a Continuing Survey of Truck and Bus Drivers

The panel contends there is a need for information on CMV drivers, their vehicles, their routes, and their employers. Given that attempts to survey CMV drivers have been infrequent and limited as to the population coverage, the panel believes that a regular survey of a sample of CMV drivers is needed to collect information on the drivers (age, gender, race), their basic health characteristics, their employment, how much they drive, their vehicles, and their routes. Also, this survey would need to be repeated on a regular basis because of the high turnover rates among the CMV driver population, as well as regular changes in the particular types and characteristics of vehicles being driven, the driving environments, the types of employment, and other variables. Repeating such a survey every 5 to 10 years would help inform the modification of guidelines for CMV drivers over time.

Since many long-haul truck drivers are not easily reached, such a

---

[3] See https://www.census.gov/svsd/www/vius/products.html [March 2016].

[4] In addition to the Vehicle Inventory and Use Survey, the Bureau of Transportation Statistics used to collect substantial data on Form M that provided details on compensation and benefits across a wide range of trucking occupations.

survey would not be easy to conduct. However, area sampling methods, such as those used in the NIOSH survey, could provide quality information. (NIOSH also used incentive payments, which would be important given that respondents would be on the clock.) While the panel is more optimistic about an area sample approach, the CDLIS could possibly serve as a sample frame for a mail or Internet survey provided that (1) it could be kept current and include information on whether the individual was employed as a truck or bus driver for greater than so many months per year, and (2) contact information could also be kept current and could include mail and Internet addresses.

The panel believes that NIOSH is best positioned to support such a data collection effort, at least in part, given the relevance of its charge to assess occupational safety and health. Because of the cost of conducting such surveys, however, other agency support undoubtedly would be needed.

> **RECOMMENDATION 1: The National Institute for Occupational Safety and Health should be enlisted to design and conduct a regularly scheduled survey every 5 to 10 years to gather information needed to better understand the demographics and employment circumstances of all commercial motor vehicle drivers in various industry segments.**

DATA AVAILABLE FROM VEHICLES

As detailed in Chapter 9, a number of new devices currently are being used or have been proposed for use with commercial motor vehicles. These devices include electronic on-board recorders that collect information on when a vehicle is in operation. Further, a number of technologies enable the collection of real-time video and telematics data, which can be used to alert a fleet manager when a vehicle is, or recently has been, speeding, hard braking, or swerving. There also are various on-board safety systems already in use or proposed for use, including indicators of weaving out of one’s lane and automatic collision avoidance systems. Such systems can provide, in real time, much of the information relevant to the driver and the vehicle listed in Table 10-1. One prime example is technology for assessing whether the driver is fatigued by measuring PERCLOS (percentage of eyelid closure). Some systems collect other, more indirect measures of fatigue, such as making quick motions with the steering wheel or multiple lane departures. In addition, information can be collected on the environment because cameras can be trained on the road ahead. When collected, however, much of this type of data is proprietary, so legal barriers currently limit researchers’ ability to acquire the data to examine the role of fatigue in highway safety. Although such data likely suffer from a lack of representativeness, the panel believes they still could play an important role in this research. For instance, they could be used to estimate upper and lower bounds for various key parameters for the entire population of CMV drivers.

It would therefore be useful to explore ways in which such data might be made available to the research community. The past 25 years has seen the growth of effective disclosure limitation techniques. Use of these techniques can greatly limit the risk of disclosure of individually identifiable information while at the same time allowing researchers to use the protected version of the data, which retains the great majority of the information. There also are arrangements whereby confidential data can be made available through research data centers such that access is allowed only for a preapproved set of researchers, who cannot take any sensitive information from the center when they have completed their analysis. Given the successful use of such techniques in many contexts, the panel believes FMCSA could benefit from working with owners of these data sets to help make them researcher-accessible and disclosure-limited. Such data sharing could be facilitated either by utilizing such techniques as cell suppression, noise addition, or production of synthetic data sets before granting researchers access or by establishing research data centers providing researchers limited access to such data for summary analyses. The following subsections describe the vehicle-based data sources currently available or on the horizon.

Data from Electronic Logging Devices

In the past few years, many carriers have decided to replace their paper logs with electronic logging devices. In the near future, all carriers may be legally obligated to do so. At a minimum, these devices record when and where the truck or bus was in operation. Therefore, in the event of a crash, the number of hours the vehicle was in operation for at least 24 hours, and likely much longer, will be automatically documented, and such data can then be used to assess whether a driver violated the HOS regulations.

It has been suggested that electronic logging devices (similar to electronic on-board recorders) provide higher-quality documentation of the number of hours a truck or bus was in operation relative to the paper logs currently used by many CMV drivers because they are more difficult to tamper with, and they document hours automatically instead of requiring the driver to do so (see also Chapter 5). The panel finds this argument very persuasive. By providing higher-quality assessments of when the truck or bus was in operation, electronic logging devices are better able to document compliance with HOS regulations, and will also provide better inputs for research examining the linkage among hours of service, fatigue, and highway safety. Accordingly, the panel believes it would be beneficial for FMCSA to compare these devices with paper logs to determine whether their use reduces HOS violations. To address the fact that those carriers with electronic logging devices may not provide representative subsets of the population of CMV drivers, this comparison could use an interrupted time series design, as described in Chapter 6. Further, the importance of switching from paper logs could be demonstrated by research showing lower crash rates for carriers that have installed these devices compared with those that have not, controlling for confounding factors with a technique such as propensity scoring.

> **RECOMMENDATION 2: The Federal Motor Carrier Safety Administration should conduct an evaluation to determine whether commercial motor vehicle drivers' use of electronic on-board recorders correlates with reduced frequency of hours-of-service violations and reduced frequency of crashes compared with those drivers who do not use such instruments.**

If either of these reductions were established, the argument for requiring the widespread use of electronic logging devices would be significantly enhanced.

A provision for Title 49 of the U.S. Code, Section 31137, mandates the use of electronic logging devices, but seems to prevent the use of the data from such devices for research, stating: "The Secretary may utilize information contained in an electronic logging device *only to enforce the Secretary's motor carrier safety and related regulations*…The Secretary shall institute appropriate measures to ensure any information collected by electronic logging devices is used by enforcement personnel *only for the purpose of determining compliance* with hours of service requirements" [emphasis added]. A recent FMCSA final rule also exists that mandates the use of electronic logging devices but could not mandate that the resulting data be used in research studies. We are therefore led to the following:

> **RECOMMENDATION 3: Given the potential research benefits of the use of data from electronic logging devices, Congress should consider modifying Title 49 of the U.S. Code to permit the use of such data for research purposes in a manner that protects individualized confidential data from disclosure, and if such a change**

is made, the Federal Motor Carrier Safety Administration should make parallel provisions in its regulations.[5]

Data Collected by Electronic Data Recorders, Vehicle Telematics, and Collision Avoidance and Fatigue Detection Systems

Many trucks and buses are equipped with electronic data recorders to record information on driver actions. Electronic data recorders are different from the electronic logging devices discussed above. They record data that are saved continuously or in response to such triggers as an acceleration exceedance (e.g., from braking, accelerating, or swerving) or a change in vehicle status (e.g., electronically sensed engine or wheel speed changes). These devices can record a wide range of data, including vehicle speed, application of brakes or clutch, steering angle, seat belt use, airbag deployment, and g-force measures associated with impact during crash sequences.

Vehicle telematics refers to various on-board technologies and wireless devices, described in Chapter 5, that transmit data to an organization in real time on how a vehicle is functioning, such as speed, acceleration and braking, airbag deployment, and crashes. They can also collect information on vehicle location. Some companies, such as Lytx, collect video data for clients, such as parents of new drivers and trucking companies, that can be used to monitor remotely how a car or truck is being driven.

The companies that currently collect such data could be encouraged to provide properly disclosure-protected data for use by researchers in examining such questions as what percentages of trip segments with and without crashes involved fatigued drivers, as measured indirectly through the operation of the vehicle or possibly more directly by adding a data capture feature for assessment of PERCLOS. In this way, research could approximate, in some sense, a continuous version of the Large Truck Crash Causation Study.

Newer devices that warn about such behaviors as lane departures (e.g., Iteris AutoVue®), driving too close to the vehicle ahead (collision avoidance systems), or abrupt steering motions have been promoted for use in alerting distracted or fatigued drivers (see also Chapter 9). While the data such devices collect depends on the specific system and are proprietary, it is reasonable to believe that they are similar to those collected by electronic logging devices and telematics systems, and as such are likely relevant to research assessing the increased crash risk associated with fatigued driving. It should be noted that drivers' acceptance of any

---

[5] A change has been made from the prepublication copy to update language to make it clear that FMCSA cannot change the law but it can modify its regulations.

system that is recording their performance or their personal information needs to be addressed.

Data Collected Prior to Serious Crashes

Black boxes on airplanes can provide detailed information on the sequence of events leading to a crash, which can then be used to suggest design modifications or implementation changes to reduce the chances of a repeat of that event. Since many of today's trucks and buses are equipped with electronic devices that record information on driver actions preceding a crash, such data (especially video data on driver physiognomy) could help in determining what factors contributed to a crash, including those related to driver fatigue or distraction. Consequently, such data would be valuable to investigators and to researchers examining relationships between driver fatigue and highway safety.

> **RECOMMENDATION 4: When commercial trucks and buses containing electronic data recorders that record data on the functioning of the driver and the truck or bus are involved in serious crashes, the relevant data should be made available to investigators and to safety researchers.**

OTHER DATA SOURCES

Other data sources that could be tapped for research on CMV driver fatigue and highway safety include vehicle inspection reports, data from large carriers, and American Transportation Research Institute (ATRI) data.

Vehicle Inspection Reports

While no technology is involved, a safety inspection system inspects 4 million commercial motor vehicles each year in North America to ensure that trucks and buses are operating safely. These inspection data are a component of the Motor Carrier Management Information System (MCMIS), compiled, maintained, and funded by FMCSA (see Chapter 5). Trained inspectors in each state examine vehicles based on criteria developed by the Commercial Vehicle Safety Alliance (CVSA). CVSA inspectors examine drivers and trucks for unsafe driving practices, HOS compliance, driver fitness, use of controlled substances, vehicle maintenance, hazardous materials compliance, and crash indications. As part of the most comprehensive Level I inspection, a driver's certificate from his or her medical examiner is checked, as is the driver's record of duty status and adherence to HOS regulations. Relevant to the present study, drivers are checked for visible signs of fatigue (although it is easy to imagine that this process misses a large fraction of cases).

In the context of this study, the utility of the data in such reports is limited because there is no direct measurement of fatigue for drivers of commercial trucks and buses. Also, the criteria used to select drivers and vehicles to inspect may differ from state to state. However, it still may be profitable to study such data on the inspection and crash history of drivers. For instance, one could determine which types of drivers, driving for what kinds of carriers, with what types of logging devices are detected by which inspection criteria to have more or fewer HOS violations. This information would have the usual deficiencies of data that are not controlled for confounding factors but might be useful for generating hypotheses.

Data from Large Carriers

Most carriers collect information on their truck drivers for such purposes as compensation, management, supervision, and monitoring. Many large carriers collect additional information on various aspects of drivers' health, their crash rates, their schedules, their routes, and their vehicles. They collect such data for various reasons—especially economic because, beyond concern about the health of drivers, drivers who frequently are involved in crashes are costly for their employers.

Driver health data can include medical information from company-sponsored clinic visits and screening exams, as well as treatment information for a variety of medical conditions, including obstructive sleep apnea (see Chapter 8). Some of the largest carriers use such data to assess the effectiveness of their health and wellness programs and to assess the economics of their insurance coverage plans.

Schneider National, for example, collects electronic log data from many trucks in its fleet that can be used to track work shift variability and the number of days since a truck was last at the home terminal. Data on such critical events as hard braking, roll stability control activation, and collision mitigation also are collected. Data collected on collisions include the time of day the crash occurred, the number of hours since the last break for the driver, and the location of the crash and the roadway type at that location. Some large carriers have relatively sophisticated data collection systems for loss events. Their loss files may include incidents that are not police-reportable or MCMIS-reportable traffic crashes. Carrier-based data also can include the history of noncrash driving, thereby addressing the lack of exposure data that exists for all trucks collectively. Crash data can be linked to personnel/work records, as well as to equipment manifests.

Clearly, such carrier-collected data could offer a rich opportunity for analysis of various questions of interest concerning HOS regulations, fatigue, and crash frequency. If data from a number of large carriers across the commercial trucking industry could be collected, organized in a database, and made available to researchers, these data could represent an important segment of the trucking industry. Such a database also would supply important information on topics on which little is otherwise known. However, it must be acknowledged that such data would exclude a large fraction of the trucking population, especially independent owner-operators, so the results of this research would not be fully generalizable.

American Transportation Research Institute Data

ATRI collects truck data through collaboration with participating carrier fleets. The ATRI database currently contains billions of data points from electronic on-board records for several hundred thousand unique vehicles spanning more than 7 years. These data, which include time, location, speed, and anonymous unique identification information, are used by ATRI researchers to produce various indicators on truck movement, highway bottlenecks, crossing times and delays, demand for truck routes, and facilities on highways. Knowing the location of a truck or bus prior to a crash could make it possible to detect whether the driver violated HOS regulations. However, this database is not available to researchers, and the data currently are collected for only a modest fraction of all trucks and buses in the United States. Therefore, these data would be of limited utility for nationally comprehensive surveillance studies. On the other hand, with such large sample sizes, the data could be useful for assessing some factors associated with an increase in crashes.

To summarize, there are or soon will be a number of data sources that could potentially provide the location of trucks and buses continuously and therefore indicate the length of time a CMV driver was driving prior to a crash. (The panel would be surprised if insurance companies did not also have relevant data.) Some direct measurement of fatigue could even be derived from assessment of PERCLOS data and indirectly from other measures. Data also will be collected on the driving environment. All of this information will exist for some subsets of the population of CMV drivers. In addition, considerable information collected on trucks and buses will provide data on crash-free driving, or exposure data. Some sources may even occasionally provide information on other characteristics listed in Table 10-1 that are needed to undertake a comprehensive assessment of the factors that cause crashes, such as health status, sleep, diet and exercise habits, number of years employed, and demographics. While these data sources will not be representative of the full population of truck and bus drivers, they still could be useful in analyzing relationships for important subsets.

A large proportion of these data is currently proprietary. Of course, even if such data could be released in a protected form, the lack of standardized fields, terminology, and definitions would be problematic. To address this problem, these elements could be standardized and disseminated, and carriers could be asked to try to match the standard format prior to data submission. Efforts to make data sets more comparable would likely be a natural by-product of the establishment of such a collective database.

In such data sets, the names of the drivers and the fleets and even the specific routes would have to be suppressed to preserve anonymity. In addition, a variety of techniques—including error inoculation, data swapping, and creation of synthetic data sets—could be applied to further protect against disclosure (see Duncan et al., 2011). Also, as discussed earlier, research data centers could provide access only to approved researchers, and review and limit the data a researcher could remove from the premises. The panel is fairly certain that such techniques could be applied successfully to data sets from these and other similar sources to permit their use for research purposes.

It must be acknowledged that there are few examples of the use of such unstructured cooperation among the private sector, nonprofits, and government agencies to substitute for what should be mainly a federal data collection program. However, the magnitude of the existing data gaps makes it unlikely that these deficiencies will be addressed in the next several years without such cooperation. The fact that relevant data are currently being collected and could be shared with researchers with almost no chance of disclosure motivates the panel's call for what is clearly a nonstandard approach to establishing a research database. In the short term, this appears to be one of the very limited possible means of acquiring this type of information that is collected by certain large carriers, as well as by ATRI.

RECOMMENDATION 5: The Federal Motor Carrier Safety Administration should incentivize those that capture driver performance data (e.g., large fleets, independent trucking associations, companies that collect telematics data, insurance companies, researchers) to increase the availability of those data relevant to research issues of operator fatigue, hours of service, and highway safety. Any such efforts should ensure that data confidentiality is maintained, per-

haps through restricted access arrangements or use of statistical techniques for disclosure protection.

SOME KEY RESEARCH QUESTIONS

The data described in the preceding sections could be used to answer many important questions related to CMV driver fatigue and highway safety. This section focuses on two specific needs for additional data collection (and analysis): (1) data on exposure and (2) driver decision making.

Exposure

One important need to support research in this area, described in greater detail in Chapter 5, is data on exposure (time spent on the road) and its relationship to other factors associated with driver fatigue and highway safety. Both exposure per hour of day—needed to compute crash rates and risk by time of day—and trip lengths and driving hours—as predictors of driver fatigue and crash risk—need to be investigated.

RECOMMENDATION 6: The Federal Motor Carrier Safety Administration should work to improve the collection of and/or access to baseline data on driving exposure by including in its data collection efforts greater detail on the driving environment and by providing these data at low levels of geographic aggregation—even for individual highway segments. Comparisons enabled by the availability of these baseline data would benefit several proposed lines of new research.

Driver Decision Making

Many factors contribute to CMV drivers' decision whether to continue driving when they recognize they are fatigued. For example, the nature of drivers' compensation likely affects their assessment of their own economic consequences (their pay) of stopping for needed rest. FMCSA is currently examining this issue in its study "Impact of Driver Compensation on Commercial Motor Vehicle Safety," a study the panel strongly supports. (The panel understands that a modification to the study design to address the possibility that the type of compensation scheme in use may be associated with other aspects of a fleet's operation is being considered, which we also support.) Factors that might be included in research on drivers' compensation are how the scheduling of work is carried out, whether time spent unloading and loading is treated as a separate component that is paid for, what time is spent waiting to load or unload, and commuting time.

Drivers' decision making—including their ability to determine whether they are too drowsy to drive safely—can be compromised if they are fatigued (see Mitler et al., 1988). Thus driver decision making could be affected by the availability and location of the nearest rest area, truck stop, or parking area and by the delivery deadline for picking up or dropping off the next load.

> **RECOMMENDATION 7: The Federal Motor Carrier Safety Administration should support research aimed at better understanding the factors associated with driver behavior related to fatigue and sleep deficiency, including what motivates drivers' decisions about whether to continue driving when they feel fatigued.**

Such research could encompass (1) barriers to healthy practices, such as compensation schemes; (2) whether management's adoption of a safety culture is beneficial in reducing fatigued driving; and (3) the impact of education and training on the degree of fatigue a driver experiences, its causes, and the possible results of fatigued driving.

METHODOLOGICAL ISSUES

As discussed in Chapter 7, data collection in studies of hours of service, fatigue, and crash risk commonly makes use of relatively standard population study designs and statistical approaches, including case-control designs, regression adjustment, and nonequivalent comparison groups. These methods can allow for the intrusion of confounding factors that are simultaneous with the intervention of interest. More generally, it is common for research on issues concerning highway safety and fatigue to be based on observational studies, which are prone to confounding influences related both to the indicator of receiving or having the "treatment" of interest and to the outcome under study (see Chapter 6). Researchers in this area, including FMCSA staff, therefore need to consider greater use of designs for nonexperimental studies that are better able to provide data that are easier to analyze in support of a clear causal statement. They also need to make use of more appropriate statistical analysis techniques that better utilize data from existing observational studies, again to counter the influence of confounding factors.

An important point emphasized in this report is that truck and bus drivers are heterogeneous with respect to the types of driving they do (see Chapter 2). Therefore, the impact of changes in HOS regulations can vary significantly for different drivers. Further, the type of driving can affect the degree to which a driver is fatigued. Thus both the nature of drivers' work-rest schedules and the type of driving they do are likely important causal factors for fatigue and for crash risk that need to be represented in any statistical model of crash risk. One way to accommodate this need is for the analysis to be stratified by employment type, but other constructs are possible as well.

It should also be noted that, as described in Chapter 6, evaluating the effects of causes rather than the causes of effects is often a more feasible and more policy-relevant goal. For example, evaluating the effect of a program designed to reduce crashes is more feasible and policy-relevant than evaluating the underlying cause of a crash (see Rosenbaum, 2002, 2009; Shadish et al., 2002). Moreover, learning about the effects of causes can help provide some understanding of the causes of effects; for example, seeing that a fatigue management program reduces collisions provides some information on the extent to which crashes are caused by fatigue.

Study Designs and Associated Data Sets

Several different types of data collection are relevant to research on driver fatigue and highway safety. Table 5-2 in Chapter 5 indicates the advantages and disadvantages of these various data sources, which vary in cost; fidelity to field operations; and ease of investigating specific driving scenarios, such as crashes. The main sources at present are crash data sets; naturalistic driving studies; simulator studies; and vehicle instrumentation, including electronic logging devices. In considering which kind of data collection to undertake to answer a specific research question, one needs to acknowledge tradeoffs in terms of control versus real-world relevance. These tradeoffs will determine whether one conducts laboratory studies and simulations, which can focus on specific situations a driver might confront, or uses field operational tests, naturalistic driving studies, and crash data sets, which can focus on how a change will be implemented in the real world and for the general population of CMV drivers. The former, more controlled experimental designs can include more (simulated) crashes and drivers with and without specific predictors. The latter, less controlled designs may offer greater fidelity but are dominated by event-free driving. In some cases, one might need to have the advantages of both to address research questions related to specific driving circumstances but applicable to the general CMV driver population. The focus here is on how designs of naturalistic driving studies, simulator studies, field experiments, and observational studies (accident reports) can be improved to better support research going forward.

Improving the Utility of Naturalistic Driving Studies

As a primary source of data on driving during noncrash periods—referred to here as exposure data—naturalistic driving studies provide an important tool for research in this field. As described in Chapter 5, such studies provide the opportunity to collect extensive data on drivers who are engaging in their typical truck and bus operations. These data on noncrash driving are extremely useful, and when crashes do occur, their cause or causes can be readily determined. Thus, naturalistic driving studies occupy an important position between crash data and simulator studies.

Several issues could be addressed to increase the utility of naturalistic driving studies. First, as with many observational studies, naturalistic driving data are collected only from volunteers, so their findings may not be generalizable to the entire population of truck and bus drivers. Second, because crashes are rare events, such studies often include fewer crashes than are needed to support estimation of exposure-outcome relationships with statistical models. This limitation motivates the use of high-kinematic events, or SCEs, as surrogate outcomes of interest. Some of these kinematic events, such as hard braking and swerving to avoid collisions, may be necessary to avoid a collision that was the fault of other drivers and may be due to a driver's alertness rather than to his or her fatigue (or distraction). Therefore, these events are not necessarily appropriate surrogate outcomes for studies on fatigued driving. (The assessment of whether an individual event is or is not related to fatigue is complicated because driving behaviors such as speed selection, lane keeping, car following, and gap acceptance all can be involved. It is similarly not clear when a crash involves driver fatigue.) Another issue is that a large amount of video data is collected, and additional research is needed on how best to identify relevant subsets of these data. The following subsections consider how these issues might be addressed.

Self-selection. To generalize any observed findings to the complete population of truck or bus drivers, one would need to collect sufficient covariate information on both the participants in a naturalistic driving study and the parent population to which one would like to generalize. Such information would have to be sufficient to enable construction of a fairly strong predictive model of which drivers would and would not volunteer for such a study. Techniques that could then be used include those based on propensity scores, poststratification, and similar methods. Some of these techniques, described in Chapter 6, require information on the parent population that does not currently exist, but could exist if the panel's Recommendation 1 on a regular sample survey were implemented.

Use of safety-critical events (SCEs). The SCEs used as surrogates for crashes in naturalistic driving studies are in theory events that in slightly different circumstances would have resulted in a crash. In practice, these events often entail the driver's causing strong g-forces on the vehicle's passengers by sudden braking or sudden veering, frequently used to avoid a collision. However, some near-crashes are not high-g-force events since, for example, another driver may have used sudden braking to prevent a crash with one of the instrumented vehicles in the study. The decision as to which SCEs to include in an analysis of a given research question and the interpretation of results using such surrogate outcomes often are difficult and subject to the criticism that the SCEs used are not representative of the same phenomenon as crashes. Guo and colleagues (2010) showed that one can obtain biased estimates of parameters when near-crash data are added to crash data.

Moreover, not all crashes should be treated equally. Some may be relatively mild bumps into a curb, while others are much more substantial and result in much greater damage and injury. There also are at-fault and not-at-fault crashes. Certainly not all crashes, whether mild or severe, are reflective of problems with driver fatigue (perhaps as many as 90% are not). Therefore, the decision as to which crashes or SCEs are indicative of problems with fatigue and should be included as outcomes and which should not be used is quite complicated and depends on the research question of interest. The crucial issue here is the need to measure outcomes that are meaningful for the research problem under investigation.

The panel advocates use of the following principles to help determine the validity of utilizing specific types of SCEs as crash surrogates. Use of SCEs is warranted, first, if they can be shown to have causal factors identical to those of crashes and, second, if there is a strong correlation in their frequency over different driving scenarios (see Guo et al., 2010). As an example, Guo and colleagues (2010) found that near-crashes provided useful information for distraction risk assessment and that use of such near-crash incidents would generate conservative estimates of risks. And Kim and colleagues (2013) found that when a g-force was beyond a prespecified threshold, it was a good predictor of crash risk.

Another reason to analyze SCEs, apart from their inclusion with crashes as outcomes of interest, is to predict future driving risk. It has been shown that high-g-force events and other SCEs are good predictors of driving risk (see Guo and Fang, 2013). Fleets (e.g., Lytx) also use telematics data on noncrashes to improve safety. Existing data sets, such as that of the Strategic Highway Research Project (SHRP) 2, can potentially be used to evaluate whether different types of SCEs are valid crash surrogates for the study of driver fatigue.

Finally, the issue of surrogate outcomes has a rich history in biostatistics and epidemiology, and recent advances in these fields could be examined for their relevance in refining understanding of the use of surrogate outcomes. (Good overviews can be found in Joffe and Green [2009] and Wittes et al. [1989]. Engineering approaches to this problem can be found in Jonasson and Rootzén [2014] and Tarko [2012].)

Use of feature extraction for video data. Research on feature extraction for video data is of continuing interest in computer science and other fields. Advances in this area could provide an alternative to kinematically defined SCEs for driving researchers. One would like to identify patterns in which behaving one way greatly lowers or raises the risk of a crash, and not behaving that way either does not affect the risk or moves it in the opposite direction. Since manually coding the continuous behavior of a driver throughout the many hours of a naturalistic driving study is extremely difficult, kinematic and similar events currently are used to identify short segments of the complete video capture for which one can code the behaviors that preceded an accident or an SCE. However, this is the same as examining crashes to assess the impact on crash risk of behaviors that preceded the crashes when one does not know how common those behaviors were during driving time when no crash or SCE occurred. It would be extremely helpful if software were developed that could alert the analyst to behaviors or situations that were good at discriminating between times of safe and unsafe driving.

Impressive research is currently being carried out on feature extraction from video. Two key research areas are understanding the degree to which a driver is fatigued from data captured by a camera focused on his or her physiognomy, and learning about the driver's external situation from a video trained on the road ahead (for details, see Hebert [2014]). Some of the challenges to such research arise because any fleet that considered installing cameras to collect such video data would likely use relatively low-cost cameras, which would not provide high enough resolution to see the driver's eyelids in all types of light or make it possible to see much more than 100 feet ahead to identify anomalous events. Therefore, it would be difficult to determine whether a driver's actions, such as swerving or braking, were warranted.

Naturalistic driving studies, such as SHRP 2, collect thousands of hours of driving data, so human assessment of such data is nearly impossible. While the time during which no crashes (or no SCEs) occur is valuable for providing baseline comparisons, the density of features of interest during noncrash time may not be very high. One therefore has at least two options. First, one can look only at interesting events that are easy to search for automatically, such as high-kinematic events and crashes. This approach can provide hypotheses for behaviors and situations that are risky. However, it cannot support assessments of increased risk since one is ignoring overall exposure data. Second, one can sample from the noncrash time periods and examine those periods for features of interest. Work by McDonald and colleagues (2013) greatly reduces the amount of video data by utilizing an alphabet indicating what actions took place. However, this data reduction approach requires that a human perform the interpretation and scoring, and until these tasks can be automated, its utility is somewhat limited. Therefore, it is difficult to determine what driver behaviors unanticipated by the researcher might be linked to increased crash risk. This is an important area for future research as cameras and pattern recognition techniques improve.

Improving Crash Reports and Data from Observational Studies in General

Most currently available data on traffic safety are collected in observational studies, which do not benefit from the relatively equal distribution of behaviors or characteristics across treatment groups that characterizes a randomized experiment. As Chakraborty and Murphy (2013) note, "In observational data associations observed in the data (e.g., between treatment and outcome) may be partially due to the unobserved or unknown reasons why individuals receive differing treatments as opposed to the effects of the treatments. Thus to conduct inference, assumptions are required." Accordingly, these data sets are limited with respect to what can be learned about the impact of a change in a regulation or the institution of a new or modified policy or program.

To address this problem, various types of observational designs can be used to enable more valid inferences about what changes might reduce crash frequency. Further, certain techniques can be used on data from observational studies to provide the balancing needed for confounders so that the comparison between the "treatment" and "control" groups approximates the comparison that would be possible with random assignment (see, e.g., Hernán and Robins, 2008). Examples of these approaches include encouragement designs, regression discontinuity, and interrupted time series designs. There are also many methods, such as difference of differences, propensity scoring, and instrumental variables, that can be used on observational data to account for unbalanced confounders. Chapter 6 introduces some of these techniques. Here the focus is on how they might be applied in the context of driving studies.

The approaches that could be attempted include both experimental and nonexperimental designs—studies in which a number of factors are randomly selected for or in which little random selection of factors is conducted, respectively. FMCSA could play a role in the design of experiments since policy interventions can be conducted in a way that facilitates randomization. For example, if FMSCA were considering changes to HOS regulations, to the North American Fatigue Management Program, or to certified medical examinations, it could first pilot the changes by rolling them out in ways that would allow for rigorous evaluation, such as by randomly selecting truck drivers to receive some new technology.[6]

There are also many scenarios in which nonexperimental designs, which often are more feasible than experiments, could be used to help isolate the causal effects of interventions. For example, an existing variation in medical examiners' policies regarding which drivers to approve could be used to look at the effects of the relevant health conditions on crash rates (under the assumption that those being evaluated were otherwise similar). This would be analogous to medical studies using physician prescribing preference as an instrumental variable for examining the effects of specific drugs on outcomes. Similarly, trends in, for example, crash rates for specific entities could be used for comparative interrupted time series designs comparing states or companies with and without policy changes.

In observational studies that compare a "treated" and an "untreated" group, one always needs to be careful about confounding due to differences between the two groups on baseline characteristics that are also associated with outcomes. For example, comparison of different types of compensation strategies could be problematic if the trucking companies that implement one compensation strategy differ on other factors (such as driver tenure, types of trucks, types of routes, and driving conditions) from those that implement the control strategy. One often can adjust for those confounding variables by using such methods as propensity scoring. Another approach to try is a comparative difference in differences analysis (a special type of interrupted time series design). (See the discussion of methods in Chapter 6.)

Perhaps more relevant to the health and wellness issues discussed in Chapter 11 are sequential multiple assignment randomized trial (SMART) designs and observational study analogues, developed by Susan Murphy and her colleagues (see, e.g., Almirall et al., 2014). Given the heterogeneous nature of truck drivers and the likelihood that they will use a variety of interventions over time, researchers in this field may find that there are applications for this new class of experimental designs.

For example, consider a naturalistic study of the effects of one versus two nights of rest after a workweek for truck drivers. A driver would

---

[6] For examples of other agencies and government groups using similar techniques, see http://www.mathematica-mpr.com/our-publications-and-findings/publications/smarter-better-faster-the-potential-for-predictive-analytics-and-rapidcycle-evaluation-to-improve [March 2016].

sometimes elect to drive after 1 day of rest and sometimes after 2 days. Here each driver is adapting the "treatment" based on his or her characteristics and previous experiences with each type of rest duration, as well as the immediately prior workweek, and these variations could be analyzed with a SMART design. The goal would be to identify optimal dynamic treatment regimes (i.e., the optimal rest patterns that adapt to previous rest). The panel believes this may be a fruitful approach for research.

Finally, as has been discussed, several different sources of data are relevant to research on driver fatigue and highway safety whose use entails trade-offs in terms of control versus real-world relevance. Each source has advantages and disadvantages with respect to support for causal inference and generalizability. Researchers could consider combining these sources to maintain their advantages and mitigate their disadvantages.

An example is the case in which policy questions can be divided into two or more components. Consider the question of whether any among several medications are fatigue inducing, such that those drivers using such medications at various dosages might be at greater crash risk. First, one might test in a simulator whether those drivers taking various doses of the different drugs had slower response times or performed worse in applying defensive driving techniques in induced crash-likely scenarios. The results would indicate which medications were worrisome and at what levels. Then, the question would be whether these findings could be generalized to the general CMV driver population and to real-world driving. To investigate that question, one might request a set of observational data from a large carrier listing the specific medications used by their drivers that had been found to be worrisome to see whether those drivers taking doses above and below the worrisome level had different crash frequencies. The simulator study would have helped reduce the chances of false positives and would have identified the medications and doses on which to focus. And to balance those drivers above and below the threshold dose for confounding factors for any given drug, techniques such as propensity scoring might be used. Much more generally, such phased approaches to research in this area might be useful.

Assessment of New Technologies for Reducing Fatigued Driving

A previous section described how electronic data recorders, driver monitoring technology, and fatigue detection technology can be used to provide relevant data for research on crash risk. Obviously, these technologies, if successful, could also accomplish their intended purpose, which is to reduce the frequency of drowsy driving and enhance driver and vehicle safety. However, the claims for new technology can sometimes exceed the realized benefits. Therefore, the panel notes here some issues that need to be addressed in the proper evaluation of such devices.

First, it is necessary that any device designed to help avoid collisions by alerting fatigued drivers be adequately field tested, including whether and how alerts are communicated to the driver and/or the fleet and the incentives, if any, for complying with the warnings. Such testing is complex and requires a human engineering approach and a human-systems integration perspective, along with expertise in the study of fatigue. This is the case because to identify effective methods for alerting drivers in a way that will get their attention, it is necessary to determine what false-negative error rates most drivers would view as tolerable, as well as other aspects of the interaction between the driver and the system.

Furthermore, the panel believes such testing would need to adhere to the following two principles: (1) tests need to be carried out by third parties to reduce the opportunity for biased assessments, and (2) care is necessary in constructing valid comparison groups, as well as in studying the contrasts between those with and without the device in question (potentially including the use of such techniques as propensity score matching). To communicate what a valid testing scheme would entail, it would be helpful for FMCSA and the National Highway Traffic Safety Administration to develop and issue a joint report indicating what they view as necessary features of an effective testing program.

> **RECOMMENDATION 8: Using a human-systems integration framework, the Federal Motor Carrier Safety Administration and the National Highway Traffic Safety Administration, in consultation with the Centers for Disease Control and Prevention and the National Institutes of Health, should develop evaluation guidelines and protocols for third-party testing, including field testing, conducted to evaluate new technologies that purport to reduce the impact of fatigue on driver safety.**

STATISTICAL ISSUES

Complex Correlation Structures

Researchers often employ standard statistical models on crash data. One can find examples of the use of logistic regression (for, say, accidents versus no accidents) or modeling of the number of accidents using Poisson regression, sometimes zero-inflated to deal with a large percentage of zero values or adjusted to accommodate overdispersion. Fatigue has transient effects on driving safety, so analysis typically needs to be conducted at a detailed level, such as the level of a trip or a driver. In this level of analysis, standard statistical models can fail to accommodate the correlation structure that is typical of data from naturalistic and other observational studies. This correlation structure is due to the fact that observations from the miles traveled by a given driver, everything else being fixed, are more highly correlated with results from other miles traveled by the same individual than with miles driven by someone else. Similarly, observations on a given segment of highway, everything else being fixed, are more highly correlated with other driving on that segment than with results from other segments. Temporal correlations may also be present, with observations at a given point in time being highly correlated with observations from nearby points in time. Such correlation structures can be handled using models with random effects for individual drivers and for road segments or through the use of models that integrate time series components. Models of this type, known as mixed-effects or hierarchical models, warrant more frequent consideration in the analysis of CMV driving data. A study by Kim and colleagues (2013) is a relevant example from the passenger vehicle literature, in which the number of high-g-force events on a trip is modeled with driver random effects and a latent temporal structure.

Another approach to handling heterogeneity in driving behavior is through the use of latent variable or hidden Markov models. These models represent heterogeneity over time within a driver through the use of assumed hidden or latent states (e.g., a high-risk driving and a low-risk driving state) whose value is inferred as part of the statistical analysis. For example, Jackson and colleagues (2015) use a two-state hidden Markov model in the analysis of trip-level driving data from a naturalistic study of teenage drivers. Methodologies drawn from survival analysis, such as recurrent event models and time-to-event models, may also be useful for the analysis of driving safety data.

Mannering and Bhat (2014) provide additional discussion of methodological approaches that have been applied to driver safety studies, together with a large number of salient references.

Power of Studies

Studies of truck and bus drivers often must make use of relatively small sample sizes, and the panel was asked to comment on the assessment of power for such studies. Before doing so, it is valuable to emphasize that the framework proposed in this chapter emphasizes attempts to estimate the magnitude of the effect of different factors on crash risk or other safety outcomes. Estimation of effect sizes, supplemented with confidence intervals that provide information about the effects of sampling variability, is a useful way to summarize the information available in a given study. Occasionally, statistical studies are summarized using a significance test; this approach is most appropriate in assessing whether a particular intervention provides any improvement over the status quo. If the goal is substantial improvement, however, significance tests alone will be insufficient. Instead, one will need to be provided with a confidence interval. Should a significance test make sense, the power of the study based on the sample size and the effect difference sought needs to be noted explicitly.

If one is using the significance of a hypothesis test as support for a research finding that a factor is (likely) causal, one must be aware of both error rates associated with the test—namely the probability of observing a significant result when that the factor has no effect, which is often set to small values such as 1 or 5 percent—and the probability of not observing a significant result when the alternative hypothesis is true—that the factor has an effect. This second error rate is one minus the power of the test. The power of a hypothesis test is sometimes ignored when a study is being designed, but a test with power of less than, say, 75 percent means that there is at least a 25-percent chance of not finding a significant result when there is one. Therefore, it is important to design a study so that the power of the alternative hypothesis is sufficiently high to trust the results. This section provides some general guidance on power analyses.

First, the alternative hypothesis often includes a range of possibilities. For example, the alternative hypothesis might be that a program had some beneficial effect on reducing driver fatigue versus the null hypothesis that the program had no or a detrimental effect. It will not be possible to obtain high power for all ranges of the alternative hypothesis since it will be difficult to distinguish a program with a very small effect from one with no effect. Furthermore, it may not be of practical importance whether a program has a very small or no effect. One needs to decide what magnitude of effect is of practical importance and choose the sample size so that the power of the study is high (e.g., greater than 80%) for all magnitudes of effects that are practically important.

In addition, calculating power post hoc is not the same as carrying out a power analysis prior to a study to determine whether it is worth conducting. This is the case because one can derive a significant result and then assess the power as being high, but that might be an artifact of a significant result that was due to chance. For details, see Hoenig and Heisey (2001).

Finally, in the case of observational studies, there is likely a bias due to the lack of balance of various confounding factors. If a sensitivity analysis can provide a bound of the impact of the confounders collectively, one would like to have power to reject the null hypothesis even allowing for bias from confounders up to the bound. (See Rosenbaum [2004], Small and Rosenbaum [2008], and Zubizarreta et al. [2013] for examples.) The panel acknowledges that the ability to do this depends on having plausible ranges of the imbalance and impact of confounders. Those assumptions can be informed by research on the confounders and their associations both with each other and with the outcomes of interest. For details, see Hsu and Small (2013) and Shepherd et al. (2007).

CONCLUDING REMARKS

As the panel has argued, research in the area of the association between CMV driver fatigue and crash risk, while often praiseworthy, has not always been reflective of current statistical methods with respect to both study design and analysis. Since the number of staff that FMCSA can devote both to writing requests for proposals and to reviewing submissions is limited, and since the number of statisticians it has on its staff is similarly small, the institution of a peer review system would help in formulating requests for proposals for research projects that would make greater use of the latest statistical methods, which, given the complex nature of data on crashes, are often an advantage. Further, such a peer review system could be used to evaluate the resulting proposals and to monitor progress after awards. Examples of agencies with such peer review systems include the U.S. Department of Education and the Agency for Healthcare Research and Quality.

FMCSA also makes use of indefinite delivery, indefinite quantity contracts (IDIQs), to facilitate contract awards. The panel believes such IDIQs need to include a broader collection of researchers with statistical expertise. Finally, for investigator-initiated studies, FMCSA needs to have more infrastructure to provide for greater interaction with such researchers while they are designing and carrying out their studies.

> **RECOMMENDATION 9: The Federal Motor Carrier Safety Administration should make greater use of independent peer review in crafting requests for proposals, assisting in decisions regarding awards, and monitoring the progress of projects (including in the study design and analysis stages). Peer review should include expertise from all relevant fields, including epidemiology and statistics—especially causal inference—to address appropriate design and analysis methods.**

11

Research Directions for Studying the Impact of Fatigue on Commercial Motor Vehicle Drivers' Health and Wellness

Truck and bus drivers are susceptible to many chronic health issues, including obstructive sleep apnea (OSA), hypertension, cardiovascular disease, adult-onset diabetes, and various other conditions commonly associated with obesity (see Chapter 2). The statement of task for this study includes the following: "The panel will also assess the relationship of these factors [hours of driving, hours on duty, and periods of rest] to drivers' health over the longer term." The need to study the relationship between fatigue and various health problems is supported in Chapter 8, where Czeisler (2015) is quoted: "Persons experiencing sleep insufficiency are more likely to have chronic diseases such as cardiovascular disease, diabetes, depression, or obesity." Chapter 8 summarizes what is currently known about the relationship between commercial motor vehicle (CMV) driving and various long-term health issues, while Chapter 9 summarizes approaches to fatigue management and health and wellness management.

This chapter describes research that the Federal Motor Carrier Safety Administration (FMCSA) and other agencies could support to address long-term health problems in the population of CMV drivers. It offers the panel's recommendations concerning fatigue management and health and wellness management, with the goal of improving the long-term health of CMV drivers. Discussed in turn are a framework for assessing factors related to CMV drivers' health and wellness; the need for an ongoing survey of CMV drivers to inform understanding of the causal role of these factors; obstructive sleep apnea (OSA) a particular fatigue-related health problem affecting drivers' health, as well as crash risk; the utility of commercial driver medical examination (CDME) data as a longitudinal health data set on CMV drivers; the need for research on drug use and driving performance; and research directions for evaluation of health and wellness programs.

A FRAMEWORK FOR ASSESSING FACTORS RELATED TO DRIVER HEALTH AND WELLNESS

Chapter 10 stresses the importance of collecting information on a wide variety of factors with a potentially important, causal role in crashes involving commercial motor vehicles, as well as on the various outcomes of interest, such as crash rate. Such information gathering is relevant here as well, to help understand the extent to which various fatigue-related risks and other causal factors impact driver health and wellness. Without such efforts, analyses may be biased by confounding influences. What is needed is a comprehensive view of what causes long-term sleep insufficiency and the long-term health conditions experienced by CMV drivers. Table 11-1 is somewhat analogous to Table 10-1 in Chapter 10, but it omits some of the rows of that table since the specific driving environment—other than what would be related to years of experience—and various factors concerning the performance of one's truck or bus would not be likely to affect a driver's long-term health profile (although this association is at least somewhat unclear since, for example, stresses due to winter or high-density driving could accumulate). The column on granularity in Table 10-1 also is omitted in Table 11-1 since most of the relevant data will be at the individual level. In addition, the outcomes of interest are much different from those in Table 10-1, focusing on a driver's long-term health in addition to long-term sleep insufficiency. Given the somewhat broader scope of Table 11-1, completing it may be more difficult than is the case for Table 10-1.

THE NEED FOR AN ONGOING SURVEY OF COMMERCIAL MOTOR VEHICLE DRIVERS

One goal is to understand the nature and extent of the various long-term health conditions experienced by CMV drivers, and so understand whether these conditions are a by-product of their occupation or would have developed regardless of the work they do. To understand what aspects of their occupation or other behaviors are causal for these conditions, it may be necessary to conduct a longitudinal survey (or a series of surveys) of CMV truck and bus drivers over time. The goal would be to collect information as their health changed over decades spent in the

TABLE 11-1 Key Domains of Factors That Influence Driver Health

Predictor Domain	Predictors/Variable Set	Database/ Data Source	Private or Public	Outcomes
Driver	• Demographics (age, race, gender) • Health conditions (high body mass index, apnea, hypertension, smoking, excessive use of alcohol) • Medications used • Frequency of acute and/or chronic fatigue (average hours driving per week, average hours on the job per week, average hours of sleep per night) • Medicine and drug use • Medical conditions • Years driving (cumulative exposure to vibration, diesel exhaust, etc.)			• Extent of long-term sleep insufficiency • Frequency of conversion from low to high blood pressure • Frequency of conversion to type II diabetes • Frequency of cardiovascular disease
Carrier	• Operation type • Fleet size • Scheduling • Logistics • Fatigue management program • Safety culture • Method of compensation			

profession. Observing what behaviors, and over what duration, are associated with changes in health status and result in these conditions may be crucial to understanding how regulatory and policy modifications to hours-of-service (HOS) regulations, requirements for medical standards and examinations, and educational programs can help reduce the risk of developing these conditions.

Specifically, to develop an in-depth understanding of the health problems of CMV drivers, one would ideally follow individual drivers for a number of years, collecting information on the nature of their job; their average number of working hours; their sleep, diet, and exercise habits; their medical condition (weight, blood pressure, use of medicines and drugs, even caffeine use); their incidence and degree of acute and chronic driver fatigue; and their crash experience—all at regular intervals. The above list may be large, but tracking many or most of these variables longitudinally would be the best way to answer key questions about CMV drivers' long-term health. To this end, previous efforts to collect fairly intrusive information from CMV drivers could be repeated and improved upon. Examples include the Commercial Driver Individual Difference Study by FMCSA, a cohort study of 20,000 drivers followed for 2-3 years, and the National Institute for Occupational Safety and Health (NIOSH) survey (Sieber et al., 2014) described in Chapter 5. The lessons learned from such studies could help in the design of the longitudinal data collection envisioned by the panel.

Clearly, the collection of information on some of these variables would have to rely on driver self-reporting, raising the possibility of misresponse. Assurance of confidentiality of any data collected would perhaps alleviate many of the concerns of the respondents. Nonetheless, questions about sleep habits and conditions related to OSA in particular are likely to elicit misresponse. If this problem became widespread, it might be necessary to conduct a medical examination for a subset of the respondents as a means of calibrating the responses.

One important objective would be to identify career drivers who do and do not develop various health conditions to determine what factors may have contributed to either a negative change in a driver's health status or maintenance of the status quo. For instance, survey data collected to date indicate that CMV drivers have a high rate of obesity, which is associated with an increased risk of OSA, diabetes, hypertension, and cardiovascular problems. To design an effective behavior modification program for reducing the frequency of obesity, one would need to know what factors help discriminate between those drivers that do and do not become obese. Another advantage of a longitudinal data collection is that it would make it possible to keep track of entry and exit to and from the profession, and as a result help in understanding the impacts of such movements on driver health and driver safety.

Given the data gaps that will be clear upon completion of Table 11-1, the panel is convinced that a longitudinal survey, or a survey administered frequently over time, of the health and wellness of CMV drivers will be necessary to collect the needed information. The establishment of a longitudinal survey has greatly improved understanding of the dynamics of changes in health characteristics in many areas of public health. Examples of such surveys include the Health and Retirement Survey and the Framingham Study. Further, the databases built with the results of these surveys have proven useful for addressing some unanticipated questions about the subject populations. Longitudinal surveys also can often support natural experiments and other types of analyses that provide clues to understanding the causal factors for various outcomes.

The costs of such an undertaking would depend on the survey's sample design, especially the sample size. The sample size would depend largely on whether subnational estimates are needed; the expected attrition rate for survey participants; the costs of following up, which are likely to be substantial given the difficulty of contacting and tracking this population; and whether the inclusion of medical measurements or tests is desired, either for the entire survey population or for a subset.[1]

While a longitudinal study design is preferable for collecting the needed information, a repeated cross-sectional design may be more feasible. A repeated cross-sectional design can provide baseline estimates and capture trends over time for variables of interest, and is appealing for studying subpopulations when sample sizes are small in individual cross-sectional data sets. Thiese and colleagues (2015b) relied on a repeated cross-sectional data set to quantify the prevalence and trends over time of multiple medical conditions in CMV drivers. The data set was drawn from the Road Ready database of CMV driver medical examinations for 2005 to 2012. There also are methods for using repeated cross-sectional studies in a way that allows inferences almost as if the studies were longitudinal. These methods include (1) designing questionnaires that ask similar questions when the survey is fielded at different periods, and (2) constructing pseudo-cohorts. One can define a cohort by restricting the time period of birth (age) or some other common characteristic (carrier size, employer, type of load). A simple example would be comparing CMV drivers aged 25 to 45 across different cross-sectional data sets. Another example would be comparing the prevalence of certain medical conditions among drivers employed by large fleets and among a group of independent owner-operators across different cross-sectional data sets.

---

[1] One way of dealing with attrition is to use refreshment samples (see Hirano et al., 2001).

However, cross-sectional studies are ultimately limited in terms of tracking changes in an individual or a cohort over time.

> **RECOMMENDATION 10: The U.S. Department of Health and Human Services and/or the U.S. Department of Transportation should fund, design, and conduct an ongoing survey that will allow longitudinal comparisons of commercial motor vehicle drivers to enable tracking of changes in their health status, and the factors likely to be associated with those changes, over time. In addition, it would be highly desirable for the survey data thus collected to include sufficient information to enable linking of the data to relevant electronic health records, with a particular focus on conditions that may threaten drivers' health and safety.**

OBSTRUCTIVE SLEEP APNEA

As noted in earlier chapters of this report, OSA is a particular problem directly associated with driver fatigue, highway safety, and driver health. As described in Chapter 8, there is very strong evidence in the case of drivers of passenger vehicles that OSA is a risk factor for negative safety outcomes (Tregear et al., 2009b), as well as for other health problems, such as hypertension. It is widely believed that a high incidence of OSA in CMV drivers also presents a significant risk of driver fatigue and therefore a safety risk on the nation's roadways.

Continuous positive airway pressure (CPAP) is the primary treatment for OSA, with an estimated 60-70 percent adherence to the therapy. Bilevel positive airway pressure or adaptive servo-ventilation is used for patients who are intolerant to CPAP. Dental devices, surgery, and weight loss are also current treatments (Jordan et al., 2014). Use of CPAP devices helps reduce the safety risk. Because of the problem of limited adherence, CMV drivers' compliance with OSA treatment protocols is likely to be a confounding factor in research addressing this issue. As a result, three groups of drivers with OSA probably need to be assessed for their long-term health and crash rates: (1) those who are compliant with their OSA treatment protocol, (2) those who are being treated but are not compliant with their protocol, and (3) those who have OSA but as yet are not being treated for it.

Obstructive Sleep Apnea Screening for CMV Drivers

FMCSA requires that CMV drivers maintain a current medical examiner's certificate to drive. CMV drivers must be examined at least every 2 years to ensure that they are fit to operate their vehicle without risk of sudden or gradual impairment or incapacitation. As described in Chapter 8, CMV drivers are to be examined by certified medical examiners who, when performing the medical exam, are to check drivers against 13 federal medical qualification standards. Of these standards, 4 are absolute and leave no discretion to the examiner other than determining whether the driver may be eligible for an exemption. For the other 9 standards, the examiner is responsible for the certification determination based on guidance issued by FMCSA.

One of the 13 medical standards states: "A person is physically qualified to drive a motor vehicle if that person has no established medical history or clinical diagnosis of a respiratory dysfunction likely to interfere with his ability to control and drive a motor vehicle safely." Even though this standard does not specifically mention OSA, OSA is cited in the advisory criteria as a respiratory condition that can interfere with oxygen exchange and pose a potential safety risk. The *Medical Examiner Handbook* contains some, although minimal, guidance on evaluation of drivers with OSA. As discussed in Chapter 8, the FMCSA Medical Review Board and a Medical Expert Panel on Obstructive Sleep Apnea and Commercial Motor Vehicle Driver Safety presented recommendations to the agency concerning screening, diagnosis, treatment, and monitoring of CMV drivers for OSA in 2008 and 2012. However, FMCSA did not adopt these recommendations. Since there is no specific guidance on criteria for evaluating drivers at risk of OSA or on treatment and follow-up, medical examiners are inconsistent in their evaluation of drivers who may be at risk of OSA.

Until September 2014, examiners could be any health care provider licensed by their state to perform physical examinations, with neither training nor certification required. The National Registry of Certified Medical Examiners (NRCME) was fully implemented only as recently as 2014. FMCSA's purpose was to have all CMV drivers examined by trained and certified medical examiners who understood the CMV driving profession and the pertinent medical standards in the agency's regulations and guidelines, including those applicable to OSA. Medical examiners may include medical doctors; doctors of osteopathy; nurse practitioners; physician assistants; and in some states chiropractors, dentists, or even physical therapists.

The absence of specific guidance to certified medical examiners on assessing CMV drivers for OSA presents challenges for employers who rely on the medical examiner to make determinations but who find that inconsistent criteria are used. FMCSA issued a bulletin to medical examiners and training associations on January 20, 2015,[2] stating that examiners

---

[2] FMCSA Bulletin to Medical Examiners and Training Organizations Regarding Obstructive Sleep Apnea. See https://nationalregistry.fmcsa.dot.gov/NRPublicUI/documents/OSA%20Bulletin%20to%20MEs%20and%20Training%20Organizations-01122015.pdf [March 2016].

should use current best practice in determining which drivers should have objective testing and offering some considerations for addressing OSA, but noting that FMCSA has no specific standards.

> **RECOMMENDATION 11: The Federal Motor Carrier Safety Administration should continue to encourage all individuals included in the National Registry of Certified Medical Examiners to utilize current best practices in identifying drivers who should be referred for additional sleep malady testing and in making determinations about commercial driver's license renewal extensions. It would be highly preferable, as soon as possible, to supply the examiners with clear criteria or guidance on when it is appropriate to refer presenting drivers for sleep malady testing.**

Need for Additional Research on Obstructive Sleep Apnea Among CMV Drivers

Understanding the linkage between OSA and crash frequency among CMV drivers is important to the charge to this panel, to FMCSA, and to the truck and bus industries. This understanding will remain incomplete if three related questions are not addressed: (1) whether OSA severity (mild, moderate, severe) influences crash risk; (2) whether OSA influences crash severity for commercial motor vehicles (fatal, injury, property damage only); and (3) whether OSA severity influences crash severity.

As noted above, evidence of an association between OSA and safety risk is strong for drivers of passenger vehicles. Mulgrew and colleagues (2008) used crash data for patients suspected to be suffering from OSA to investigate the association between OSA severity and crash severity. The study found that patients with OSA were at an increased risk of crashes and that the crash risk did vary by OSA severity (2.6 times higher for mild OSA, 1.9 for moderate, and 2.0 for severe), but did not increase monotonically as suspected. The crash risk was disproportionately higher in the case of crashes that involved a personal injury (4.8 times higher for mild OSA, 3.0 for moderate, 4.3 for severe) (Mulgrew et al., 2008).

There also remain a number of key questions concerning OSA and CMV drivers. These include (1) what percentage of CMV drivers are affected; (2) what the increased crash risk is for CMV drivers for varying degrees of apnea-hypopnea (i.e., the severity of OSA); (3) how best those at risk for OSA can be identified as a result of the observations and tests conducted by nationally registered medical examiners; (4) to what extent CPAP and related treatment technologies reduce the risk associated with OSA; and (5) what length of treatment and what degree of treatment compliance (i.e., the number of hours a night and number of nights a year CPAP is used) will generate such reductions.

Related to the question of identifying drivers who should be referred for diagnostic testing for OSA, expert panels and advisory boards have issued a number of recommendations as to characteristics that could identify which drivers should be evaluated for OSA. Those recommendations include utilizing body mass index (BMI) above a specified value, enlarged neck circumference, or obstructed posterior throat, or certain medical conditions such as hypertension. As noted above, however, while there are various potential screening criteria for OSA, FMCSA has not indicated which of these should be used in certifying medical exams, and as a result, medical examiners do not apply consistent criteria with respect to OSA.

What is needed is clear-cut guidance from FMCSA on the criteria medical examiners could/should use to determine which drivers need to be referred for diagnostic sleep disorder testing, on acceptable diagnostic criteria, on the level of apnea-hypopnea index (index of severity of OSA) that would necessitate treatment, and on the acceptable duration and frequency of treatment. Also needed is guidance on what criteria should be used for removing a driver from service by restricting his/her commercial driver's license (CDL) because of the presence of untreated OSA and for how long if indicated.

The problem of finding a statistical rule for screening a population for testing by identifying groups of people much more and much less likely to have a characteristic is a classic problem in discriminant analysis. Many statistical techniques can be used to find excellent rules given a "training set" of input data (data for a set of drivers indicating whether they have OSA and their health characteristics that would be available to a medical examiner) on drivers with varying degrees of characteristics that may play a role in determining the screening rule and with varying degrees of OSA severity. The characteristics of interest include weight, height, neck circumference, degree of obstructed posterior throat, and degree of hypertension. The techniques that can be utilized include normal theory-based discriminant analysis, logistic regression, classification trees, neural nets, and support vector machines. Given a good training set, finding a good discriminant rule can be straightforward.

Screening rules need to have low errors of two types: the rules need to rarely indicate that drivers should have a test for OSA when they do not have an extreme case of OSA and to rarely indicate that drivers do not need to be tested when they do have an extreme case. It would be difficult to know the possible levels of these two error rates prior to testing. It is relatively clear that the error of not testing when it is called for is more important than the error of testing those who turn out not to have severe cases of OSA. Thus it might be sensible to set the error of not testing when needed to some acceptable low value and then choose the procedure that minimizes the error of testing when unnecessary.

> **RECOMMENDATION 12: The Federal Motor Carrier Safety Administration should support peer-reviewed research on obstructive sleep apnea (OSA) and commercial motor vehicle drivers throughout all the research stages, from the drafting of requests for proposals through analysis of data. The supported research should be focused on a better understanding of the incidence of OSA in commercial motor vehicle drivers; its impact on driver fatigue, safety, and health; and the benefits of treatments. Specific research topics might include**
>
> - **determining the number of commercial motor vehicle drivers whose quantity/quality of sleep and driving performance are likely affected at various levels of apnea-hypopnea (index of OSA severity);**
> - **determining what rules for sleep-screening referrals are effective in discriminating between those commercial motor vehicle drivers with and without OSA;**
> - **delineating the causal chain from diagnosis of OSA (preferably as a function of severity) to increased likelihood of crash frequency among commercial motor vehicle drivers;**
> - **determining the impact of treatment with positive airway pressure (PAP) and similar devices on long-term health and crash rates among commercial motor vehicle drivers with varying degrees of apnea severity; and**
> - **identifying the required/recommended duration of initial PAP treatment (e.g., suggested number of hours of treatment per day/week) before a driver can be certified to return to driving.**

UTILITY OF COMMERCIAL DRIVER MEDICAL EXAMINATION DATA

During the medical exam required every 2 years for a CMV driver to maintain his or her CDL, the medical examiner notes the presence and absence of numerous health conditions and provides certification decisions. There are web-based platforms wherein the medical examiner can store results of the examination and certification decisions. For example, Thiese and colleagues (2015b) obtained CDME data for 88,246 CMV drivers from such a web-based platform—Road Ready, Inc. They studied medical data for the years 2005 to 2012 and analyzed them for associations among BMI, medical disorders, and driver certification (Thiese et al., 2015a).

CDME data are a valuable source of information on driver demographics, medical history, height, weight, blood pressure, heart rate, urinalysis, and medical examinations. As all CMV drivers undergo the medical examination to maintain their CDL, the CDME data capture both drivers employed by companies and independent owner-operators. Given that the medical examination is conducted at least once every 2 years, the CDME data can become a longitudinal health data set on CMV drivers, as Thiese and colleagues (2015b) demonstrated. One can use these data not only to calculate baseline estimates and trace the prevalence of various health conditions in CMV drivers (Thiese et al., 2015b) but also to estimate the impact of new and revised guidance on disqualifying medical conditions for driver certification.

THE NEED FOR RESEARCH ON DRUG USE AND DRIVING PERFORMANCE

One of the findings of the Large Truck Crash Causation Study was that 17 percent of truck drivers in the study sample were using over-the-counter drugs, and 2 percent were using illegal drugs. More than 27 years ago, Lund and colleagues (1988) conducted a health survey at a truck weighing station in Tennessee involving 317 randomly selected tractor-trailer drivers. Participating drivers were asked to provide urine or blood samples, which were screened for alcohol and 80 other substances. Twenty-nine percent of the drivers in the study sample had alcohol, marijuana, cocaine, or prescription or nonprescription stimulants in their blood and urine. Couper and colleagues (2002) found similar results when they investigated the prevalence of drug use among 1,067 drivers in the state of Oregon. Twenty-one percent of the urine specimens tested positive for illicit, prescription, and/or over-the-counter drugs, and 7 percent tested positive for more than one drug. In addition, there have been numerous international studies of drug use among CMV drivers, as well as attempts to assess the performance effects. For details, see Krueger et al. (2011), which includes an extensive list of salient references.

The issue of particular concern to this panel is how various drugs influence the driving performance of CMV drivers. The U.S. Food and Drug Administration (FDA) and other federal agencies have advised the public that some prescription and over-the-counter medicines can make it unsafe to drive because their use may cause drowsiness (see Krueger et al., 2011). Research is scant on the link between drug use and impairment among CMV drivers as both ethical and practical difficulties are entailed in learning about drug impairment from actual motor vehicle accidents.[3]

---

[3] Presentation by Ronald Farkas, FDA, at National Transportation Safety Board (NTSB) Drowsy Driving Forum, NTSB Conference Center, Washington, D.C., October 24, 2014.

Nevertheless, some work has been done to investigate this linkage. The Large Truck Crash Causation Study found that prescription drug use did not increase the risk of being involved in a crash, but the relative risk of over-the-counter drug usage and illegal drug usage was 1.3 and 1.8 times higher, respectively. The National Transportation Safety Board, in collaboration with the National Institute on Drug Abuse, conducted drug screens on blood specimens of 168 fatally injured truck drivers to investigate the influence of alcohol and other drug usage on fatal-to-the-driver crashes. The study found concentrations of marijuana and alcohol that could lead to driver impairment (Crouch et al., 1993).

To address the issue of drugs and driving impairment would require information on the prevalence of drug use (including details on the types of substances) among CMV drivers, which could be one of the data items in the longitudinal survey recommended above. Also needed is comprehensive research investigating the association of alcohol and other substances with driving performance. For a summary of what is known, see Krueger (2010a).

RESEARCH DIRECTIONS FOR EVALUATION OF HEALTH AND WELLNESS PROGRAMS

From 1996 to 2006, FMCSA and the American Transportation Research Institute (ATRI), the research arm of the American Trucking Associations (ATA), conducted an educational program on fatigue for trainers of motor carriers, their trucking officials, and drivers. This program—a train-the-trainer program entitled Mastering Alertness and Managing Driver Fatigue—covered such topics as the importance of obtaining adequate rest and sleep, body and sleep physiology, circadian rhythm effects, shift-lag influences from rotating work schedules, sleep disorders, the influences of chemical substances, a list of drowsy driver warning signals, and a set of fatigue countermeasures (Krueger et al., 2007). This program was accompanied by a driver wellness train-the-trainer program known as "Gettin'-In-Gear," offered by FMCSA and ATRI from 2001 to 2006. The latter program focused on the health, fitness, and wellness of CMV drivers and covered such topics as various health conditions affecting these drivers, sleep disorders, drug and alcohol use, individual diet and exercise plans to improve health and wellness, and relaxation techniques (Krueger and Brewster, 2002).

Starting in July 2013, FMCSA and its international partners in Canada began offering the North American Fatigue Management Program (NAFMP) (see Chapter 8). However, FMCSA does not know the extent to which these PowerPoint slides are being read by CMV drivers or the extent to which reading them is helping to change the drivers' behaviors to reduce their susceptibility to fatigue. (The primary program evaluation activities can be found in Moscovitch et al. [2006] and Smiley et al. [2009].) Answering either of these questions will not be easy. Any type of request for personal information from visitors to the website is likely to have a very high percentages of nonresponse, and attempts to collect such information may reduce the number of visitors. Absent any type of evaluation, however, FMCSA does not know the extent to which the current program is working. The agency has held at least one set of preliminary user focus group interviews, but has not yet expressed a clear idea as to which of the 10 training modules might be in need of modification. More generally, FMCSA needs to include in any education or training program a summative evaluation phase to test its efficacy

There are two possible approaches FMCSA could use to evaluate the NAFMP. First, a surprising amount of information can be acquired passively through web analytics. This information includes (1) the number of visits to the website; (2) the number of unique visitors; (3) the number of page views per visit; (4) the average visit duration; (5) the percentage of people who visit the site and immediately move on without looking at any other pages; and (6) the number of first-time visitors, which can be used to determine the percentage of new versus returning users. One can also learn about the different devices used to access the site, which tells something about the visitors. Further, one can learn whether a visitor was searching for a particular type of content, was referred to the site by another site, or typed in the website's address directly. One also can identify the variety of keywords that brought a user to the website from a search engine. These various statistics can be cross-classified to obtain greater detail, and these statistics are available for 30 days to enable examining trends over time. To FMCSA's credit, much of this type of analysis can be found in the monthly reports prepared by ATRI for the NAFMP Steering Committee. The panel supports the continuation of these analyses. Additionally, FMCSA needs to find a way of identifying when NAFMP training modules are downloaded from the website for corporate use, and determining whether these training materials were subsequently used in group training, say, at a carrier's own training classroom equipped to reach larger numbers of drivers.

Analyses of individual interactions with the NAFMP website are limited because one cannot determine whether a visitor has changed his or her behavior after completing various course modules and whether those who have made such recommended lifestyle changes have actually improved their health. For FMCSA to learn about the effectiveness of the NAFMP, it will be necessary to recruit a sample of truck and bus drivers to learn what their interaction with the program has been, whether this degree of interaction has affected their behavior, and whether there has been an associated change in their degree of fatigue or their health status. Clearly, if such a questionnaire were directed at web visitors, one would have to rely on self-reports to determine changes in health status, which, as discussed above, could be subject to considerable misresponse. Instead, direct measures of, for example, the amount of sleep obtained, current blood pressure, and weight would be desirable and perhaps even necessary. As a stand-alone survey, this inquiry would likely be costly. However, the collection of such information could be incorporated into the longitudinal data collection mentioned in Recommendation 10, further justifying that recommendation.

In addition to the NAFMP, some success in modifying behavior has recently been achieved through the use of incentive-based programs, such as the Safety and Health Involvement for Truckers Program, funded by the National Institutes of Health's National Heart, Lung, and Blood Institute and targeted at truck drivers aiming to manage or lose weight.[4] Such a program could have important advantages over the NAFMP in modifying behavior. Therefore, research is needed to examine the advantages and disadvantages of such an approach.

RECOMMENDATION 13: The Federal Motor Carrier Safety Administration (FMCSA) should carry out a research program on driver fatigue management and training. This research program should include

- **evaluating the effectiveness of the North American Fatigue Management Program (NAFMP) for educating truck and bus drivers in how to modify their behavior to remedy various potential sources of fatigue;**
- **determining how effective the NAFMP training modules are in meeting the needs of drivers' employers, including fleet managers, safety and risk managers, dispatchers, driver trainers and other corporate officials (e.g., those conducting carrier-sponsored employee health and wellness programs);**
- **evaluating any new education programs regarding sleep apnea that FMCSA has or plans to develop; and**
- **examining possibilities for the development and evaluation of incentive-based programs for improving health and fitness, including regular coaching, assessment, and support.**

---

[4] Safety & Health Involvement for Truckers. Available: https://www.ohsushift.com/ [March 2016].

Glossary

Actigraphy: The monitoring of time spent asleep using an actigraph, a device often worn on the wrist that assumes lack of movement indicates the wearer is asleep.

American Transportation Research Institute (ATRI): A nonprofit research organization that is a member of the American Trucking Associations.

Biomathematical models: Mathematical models that predict the impacts on performance of various work/rest schedules, including the effects of extended wake durations or rotating shifts. Very simple versions of such models are used to analyze two systems: sleep/wake homeostasis and the circadian biological clock.

Body mass index (BMI): A measure of a person's degree of obesity, which is obtained by dividing the person's weight (in kilograms) by the square of his or her height (in meters).

Carrier: A truck or bus carrier is a business that owns trucks or buses, respectively, and, unlike independent owner-operators, employs drivers to meet its driving needs.

Case-control study, cohort study, case-crossover study: A *case-control study* compares subjects who have a response or outcome of interest with those who do not in order to determine whether the two populations differ in the frequency of potential causal factors. This analysis assumes that the two groups of subjects are otherwise similar with respect to any confounding factors, and this similarity can be achieved through pairwise matching and other techniques. In a *cohort study*, subjects are followed prospectively to determine which risk factors are and are not associated with the development of some condition or outcome of interest over time. A *case-crossover study* is similar to a matched-pair case-control design, except that the matched pair in this case is the same subject at two different points in time.

Circadian rhythm: Based on an internal biological clock that regulates when one has periods of sleepiness and wakefulness during the day. For many people, the period of greatest sleepiness comes between 2:00 and 4:00 AM, with a lesser period occurring between 2:00 and 4:00 PM.

Commercial driver's license (CDL): Required to drive a commercial motor vehicle. To obtain a CDL, one must pass both a skills test and a knowledge test.

Commercial Driver's License Information System (CDLIS): A nationwide computer system of commercial driver's licenses that allows state driver's licensing agencies to determine whether a commercial driver has any out-of-state convictions or other similar information. The central site is maintained by the American Association of Motor Vehicle Administrators.

Commercial motor vehicle (CMV): Either a single vehicle with a gross vehicle weight rating of 26,001 or more pounds; a combination of vehicles with a gross vehicle weight rating of 26,001 or more pounds if a vehicle being towed is more than 10,000 pounds; a vehicle that carries 16 or more passengers, including the driver; or a vehicle that transports hazardous materials.

Commercial Vehicle Safety Alliance (CVSA): An international nonprofit organization comprising motor carrier safety officials and industry representatives from the United States, Canada, and Mexico, with the mission of promoting commercial motor vehicle safety and security. As part of its mission, CVSA establishes out-of-service standards for commercial vehicles operating in North America, and its inspections can place a vehicle out of service if violations are discovered during the inspection process.

Confounding factor: A factor that may be causally associated with some outcome of interest and is not of primary interest in a study.

Disclosure avoidance, disclosure protection (sometimes called statistical disclosure avoidance, statistical disclosure protection, disclosure control, or other similar terms): A collection of techniques used to protect confidential individual-level information while at the same time retaining as much of the information as possible in a database that does not provide individually identifiable information.

Drowsiness or fatigue: *Drowsiness* refers to feeling sleepy or tired or being unable to keep one's eyes open. *Fatigue* is a more general and subjective term that refers to increasing performance variability and instability in behavioral alertness and vigilance due to continued time on task without breaks.

Electronic-on-board recorder (EOBR): A device that is used primarily to log when a vehicle was in operation for some recent period of time—essentially the same as electronic logging devices (ELDs), and in contrast to paper logs.

Employee health and wellness programs: Encompass various methods for educating, incentivizing, and providing feedback to employees about their diet, exercise, and sleep habits to improve their health and wellness over time.

Exposure data: The amount of time or distance driven. Vehicle-miles traveled (VMT) is a common measure of exposure.

Fatality Analysis Reporting System (FARS): A database containing information on all motor vehicles involved in fatal traffic crashes in the United States, maintained by the National Highway Traffic Safety Administration (NHTSA).

Fatigue risk management plan (FRMP) and fatigue risk management system (FRMS): A *fatigue risk management plan* establishes policies on managing and mitigating fatigue during operations. It typically includes a requirement for employee (e.g., drivers, fleet managers, dispatchers) fatigue awareness training, as well as processes for reporting instances of fatigued driving. A *fatigue risk management system* manages operator fatigue at a more granular level, and includes a continuous feedback loop that provides a means for continuous measurement and monitoring of an individual worker's schedules.

Federal Highway Administration (FHWA): An agency within the U.S. Department of Transportation that, through financial and technical assistance to state and local governments, supports the design, construction, and maintenance of the U.S. highway system so that it is safe and technologically up to date.

Federal Motor Carrier Safety Administration (FMCSA): An agency within the U.S. Department of Transportation with the mission of preventing commercial motor vehicle-related fatalities and injuries. Its activities include enforcing safety regulations, targeting high-risk carriers and commercial motor vehicle drivers, improving safety information systems and commercial motor vehicle technologies, strengthening commercial motor vehicle equipment and operating standards, and increasing safety awareness.

General Estimates System (GES): A database based on a hierarchical stratified sample of police-reported crashes involving at least one motor vehicle engaged in travel on a roadway and resulting in property damage, injury, or death, maintained by the National Highway Traffic Safety Administration (NHTSA).

Hours-of-service (HOS) regulations: Set by the Federal Motor Carrier Safety Administration (FMCSA), they specify the maximum number of hours in a day and in a work week that commercial motor vehicle drivers can drive and work, along with other rules on breaks and restart provisions.

Hypopnea/apnea: Measures of a lack of oxygen during sleep. The number of times a person is awakened during sleep per hour is called the apnea/hypopnea index. It includes breathing cessations, or apneas, and partial obstructions, or hypopneas.

Instrumental variables: A statistical technique that controls for confounding and measurement error in observational studies and so promotes the potential for drawing causal inferences. In particular, in regression models, an explanatory variable X may be associated with the error term (which combines the effects of factors not included in the model), with the result being biased regression coefficients. An instrument Z is one that is causally related to X, that affects the outcome variable Y only through its impact on X, and that is independent of the error term, and therefore in replacing X, will provide unbiased estimates of the regression coefficients.

Large Truck Crash Causation Study (LTCCS): A collaborative research project between the Federal Motor Carrier Safety Administration (FMCSA) and the National Highway Traffic Safety Administration (NHTSA) to collect detailed information on the possible causes of 963 large-truck crashes that took place between 2001 and 2003. The resulting database has supported a great deal of study of the causes of large-truck crashes.

Motor Carrier Management Information System (MCMIS): A census of all trucks and buses involved in a crash that involved a fatality, an injury to a person transported for immediate medical attention, or at least one vehicle towed because of disabling damage. It is maintained by the Federal Motor Carrier Safety Administration (FMCSA).

National Highway Traffic Safety Administration (NHTSA): Established to prevent highway crashes, in part by supporting research on their causes and how they can be prevented. Its focus is more on the safety of the vehicle and the driving environment than on the driver.

National Institute for Occupational Safety and Health (NIOSH): An agency within the Centers for Disease Control and Prevention that carries out research to reduce work-related illnesses and injuries.

National Registry of Certified Medical Examiners (NRCME): A list, established by the Federal Motor Carrier Safety Administration (FMCSA), of medical professionals who have completed training and successfully passed a test on FMCSA's physical qualification standards for working as a commercial motor vehicle driver. These professionals can then be used to determine whether a commercial motor vehicle driver is physically fit to drive. A commercial motor vehicle driver whose medical certificate expires must be examined by a medical professional listed in the NRCME to retain driving privileges.

National Transportation Safety Board (NTSB): Charged with determining the probable cause of transportation accidents and promoting transportation safety, and with assisting victims of transportation accidents and their families.

Naturalistic driving study (NDS): A study that collects video and other data on the performance of a vehicle while the driver is carrying out his or her usual driving duties. The Strategic Highway Research Program (SHRP) 2 is a recently completed large naturalistic driving study of automobile drivers.

North American Fatigue Management Program (NAFMP): An online educational program designed to promote greater understanding of effective ways to manage and mitigate fatigue in trucking and busing operations.

Observational study: A study in which researchers observe subjects and measure variables of interest without assigning treatments to the subjects, so that the treatment each subject receives is beyond the control of the researcher.

Obstructive sleep apnea (OSA): The most common type of sleep apnea, a sleep disorder in which breathing repeatedly stops and starts during sleep. In OSA, one's throat muscles intermittently relax and block one's airway.

Odds ratio: A measure of the association between a risk factor and an outcome. The odds ratio is the probability that an outcome will occur divided by the probability that it will not in the presence of a risk factor, divided by the same ratio when the risk factor is not present.

Owner-Operator Independent Drivers Association (OOIDA): A North American trade organization that represents the interests of truck drivers and works to affect state and federal legislation regarding the trucking industry.

PERCLOS (percentage of eye closure): A measure of the percentage of eyelid closure over a period of time—one of the most accepted measures of drowsiness.

Polysomnography: A sleep study that records the following body functions as subjects either sleep or try to sleep: air flow in and out of the lungs, the level of oxygen in the blood, body position, brain waves, breathing effort and rate, electrical activity of muscles, eye movement, and heart rate.

Positive airway pressure (PAP)/continuous positive airway pressure (CPAP) devices: Devices that are effective for treating obstructive sleep apnea. They work by blowing pressurized air through the airway to keep the throat open.

Propensity scores: A statistical technique that uses an estimate of the probability of treatment as a function of a set of potential confounding factors in various ways to balance a study's treatment and control groups for the effect of those confounding factors.

Psychomotor vigilance test (PVT): A reaction-timed test that measures behavioral alertness and attention by measuring psychomotor speed, lapses of attention, and impulsivity induced by fatigue.

Restart provision: As part of the current hours-of-service regulations, drivers may "restart" their 7/8-consecutive-day duty period after 34 or more consecutive hours off duty.

Rubin's causal model: A causal effect is defined as the nonzero difference between potential outcomes (one of which must be counterfactual) when different treatments are administered at a particular point in time under otherwise identical conditions.

Rumble strips: A road safety feature designed to alert fatigued or inattentive drivers that they are leaving their lane or the roadway by causing a vibration and an audible rumbling sound when a vehicle's tires pass over them.

Safety-critical events (SCEs): Often identified by various kinematic motions of the vehicle, they are thought to be events that could have been crashes if the circumstances of the driving environment had been only slightly different.

Safety culture: A culture of shared beliefs, practices, and attitudes within an establishment that shapes behavior. It can include management and employee norms, beliefs, and attitudes; policies and procedures; supervisor priorities; responsibilities and accountability; and employee training and motivation.

Short-haul vs. long-haul drivers: Short-haul drivers often drive no farther than 150 miles from home and return home most evenings. Long-haul drivers often drive farther than 150 miles, and their trips often require them to sleep away from home.

Technologies for crash avoidance: Include electronic stability control (ESC), roll stability control (RST), lane departure warning (LDW), blind spot warning (BSW), forward collision warning (FCW), adaptive cruise control (ACC), and collision mitigation braking systems (CMBS).

Telematics: A set of technologies that use sensors to monitor the performance of a motor vehicle, including braking, steering, and speeding.

Truckload versus less-than-truckload services: Less-than-truckload services combine shipments from multiple shippers and transports to multiple destinations. Truckload services typically entail a load from a single shipper that is to be transported to a single destination.

Trucks Involved in Fatal Accidents (TIFA) and Buses Involved in Fatal Accidents (BIFA): Two databases that resulted from censuses of medium and heavy trucks and of buses, respectively, involved in fatal crashes. Both were based on the Fatality Analysis Reporting System (FARS), supplemented by data collected by University of Michigan Transportation Research Institute (UMTRI) researchers.

Vigilance: "The ability to maintain sustained attention within the road environment" (Thiffault and Bergeron, 2003). *Alertness* is roughly synonymous with vigilance.

References

Abe, T., Nonomura, T., Komada, Y., Asaoka, S., Sasai, T., Ueno, A., and Inoue, Y. (2011). Detecting deteriorated vigilance using percentage of eyelid closure time during behavioral maintenance of wakefulness tests. *International Journal of Psychophysiology, 82*(3), 269-274.

Abe, T., Komada, Y., and Inoue, Y. (2012). Short sleep duration, snoring and subjective sleep insufficiency are independent factors associated with both falling asleep and feeling sleepiness while driving. *Internal Medicine*, *51*(23), 3253-3260.

Abe, T., Goel, N., Basner, M., Mollicone, D., Rao, H., and Dinges, D.F. (2015). Integration of sleep need and fatigue mitigation into human system operation. In F. Durso and D. Boehm-Davis (Eds.), *Handbook of Human Systems Integration* (pp. 177-191). Washington, DC: American Psychological Association.

Akerstedt, T., Connor, J., Gray, A., and Kecklund, G. (2008). Predicting road crashes from a mathematical model of alertness regulation—The Sleep/Wake Predictor. *Accident Analysis & Prevention, 40*(4), 1480-1485.

Aldrich, M.S. (1989). Automobile accidents in patients with sleep disorders. *Sleep*, *12*(6), 487-494.

Almirall, D., Griffin, B.A., McCaffrey, D.F., Ramchand, R., Yuen, R.A., and Murphy, S. (2014). Time-varying effect moderation using the structural nested mean model: Estimation using inverse-weighted regression with residuals. *Statistics in Medicine, 33*(20), 3466-3487.

American Transportation Research Institute. (2010). *Hours-of-Service Rules Safety Impacts, 2010 Analysis*. Arlington, VA: American Transportation Research Institute. Available: http://www.atri-online.org/ATRI_HOS_Analysis_2010.pdf [March 2016].

Ancoli-Israel, S., Czeisler, C.A., George, C.F.P., Guilleminault, C., and Pack, A.I. (2008). *Expert Panel Recommendations: Obstructive Sleep Apnea and Commercial Motor Vehicle Driver Safety*. Available: https://www.fmcsa.dot.gov/sites/fmcsa.dot.gov/files/docs/Sleep-MEP-Panel-Recommendations-508.pdf [March 2016].

Angrist, J.D., Imbens, G.W., and Rubin, D.B. (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association, 91*(434), 444-455.

Antin, J., Stulce, K., Eichelberger, L., and Hankey, J. (2015). *Naturalistic Driving Study: Descriptive Comparison of the Study Sample with National Data* (SHRP2 Report S2-S31-RW-1). Available: http://onlinepubs.trb.org/onlinepubs/shrp2/SHRP2_S2-S31-RW-1.pdf [March 2016].

Anund, A., Kecklund, G., Peters, B., Forsman, Å., Lowden, A., and Åkerstedt, T. (2008). Driver impairment at night and its relation to physiological sleepiness. *Scandinavian Journal of Work, Environment and Health*, *34*(2), 142-150.

Apostolopoulos, Y., Sonmez, S., Shattell, M.M., and Belzer, M. (2010). Worksite-induced morbidities among truck drivers in the United States. *AAOHN Journal: Official Journal of the American Association of Occupational Health Nurses*, *58*(7), 285-296.

Baiocchi, M., Small, D., Lorch, S., and Rosenbaum, P. (2010). Building a stronger instrument in an observational study of perinatal care for premature infants. *Journal of the American Statistical Association*, *105*(492), 1285-1296.

Balkin, T., Thorne, D., Sing, H., Thomas, M., Redmond, D., Wesensten, N., Williams, J., Hall, S., and Belenky, G. (2000). *Effects of Sleep Schedules on Commercial Motor Vehicle Driver Performance* (DOT-MC-00-133). Washington, DC: U.S. Department of Transportation, Federal Motor Carrier Safety Administration.

Banerjee, I., Lee, J., Jang, K., Pande, S., and Ragland, D. (2009). *Rest Areas: Reducing Accidents Involving Driver Fatigue*. Berkeley: University of California Traffic Safety Center, California Department of Transportation.

Banks, S., and Dinges, D.F. (2007). Behavioral and physiological consequences of sleep restriction. *Journal of Clinical Sleep Medicine*, *3*(5), 519-528.

Banks, S., Van Dongen, H.P.A., Maislin, G., and Dinges, D.F. (2010). Neurobehavioral dynamics following chronic sleep restriction: Dose-response effects of one night for recovery. *Sleep*, *33*(8), 1013-1026.

Barbe, F., Pericas J., Munoz, A., Findley, L., Anto, J.M., and Agusti, A.G. (1998). Automobile accidents in patients with sleep apnea syndrome. An epidemiological and mechanistic study. *American Journal of Respiratory and Critical Care Medicine*, *158*(1), 18-22.

Barr, L.C., Yang, C.Y.D., Hanowski, R.J., and Olson, R. (2011). *An Assessment of Driver Drowsiness, Distraction, and Performance in a Naturalistic Setting* (FMCSA-RRR-11-010). Washington, DC: U.S. Department of Transportation, Federal Motor Carrier Safety Administration.

Basner, M., and Dinges, D.F. (2009). Dubious bargain: Trading sleep for Leno and Letterman. *Sleep*, *32*(6), 747-752.

Basner, M., and Dinges, D.F. (2011). Maximizing sensitivity of the Psychomotor Vigilance Test (PVT) to sleep loss. *Sleep*, *34*(5), 581-591.

Basner, M., and Dinges, D.F. (2012). An adaptive-duration version of the PVT accurately tracks changes in psychomotor vigilance induced by sleep restriction. *Sleep*, *35*(2), 193-202.

Basner, M., Fomberstein, K., Razavi, F.M., Banks, S., William, J., Rosa, R., and Dinges, D.F. (2007). American Time Use Survey: Sleep time and its relationship to waking activities. *Sleep*, *30*(9), 1085-1095.

Basner, M., Mollicone, D., and Dinges, D.F. (2011). Validity and sensitivity of a brief Psychomotor Vigilance Test (PVT-B) to total and partial sleep deprivation. *Acta Astronautica,* 69(11-12), 949-959.

Basner, M., Spaeth, A.M., and Dinges, D.F. (2014). Sociodemographic characteristics and waking activities and their role in the timing and duration of sleep. *Sleep*, *37*(12), 1889-1906.

Battelle-Seattle Research Center. (2000). *Stress and Fatigue Effects of Driving Longer Combination Vehicles*. Washington, DC: U.S. Department of Transportation, Federal Motor Carrier Safety Administration.

Bedard, M.A., Montplaisir, J., Malo, J., Richer, F., and Roulean, I. (1993). Persistent neuropsychological deficits and vigilance impairment in sleep apnea syndrome after treatment with continuous positive airways pressure (CPAP). *Journal of Clinical Experimental Neuropsychology*, *15*(2), 330-341.

Beilock, R. (1995). Schedule-induced hours-of-service and speed limit violations among tractor-trailer drivers. *Accident Analysis and Prevention*, *27*(1), 33-42.

Belenky, G., Wesensten, N.J., Thorne, D.R., Thomas, M.L., Sing, H.C., Redmond, D.P., Russo, M.B., and Balkin, T.J. (2003). Patterns of performance degradation and restoration during sleep restriction and subsequent recovery: a sleep dose-response study. *Journal of Sleep Research, 12*, 1-12.

Belenky, G., Wu, L.J., Zaslona, J.L., and Hodges, J. (2012). *Motorcoach Driver Fatigue Study, 2011* (FMCSA-RRR-12-042). Washington DC: U.S. Department of Transportation, Federal Motor Carrier Safety Administration.

Belenky, G., Hanowski, R., and Jovanis, P. (2013). *Expert Panel Report: Fatigue and Commercial Motorcoach/Bus Driver Safety*. Presented to the Federal Motor Carrier Safety Administration. Available: https://www.fmcsa.dot.gov/sites/fmcsa.dot.gov/files/docs/Expert%20Panel%20Recommendations%20-%2012-20-2012%20EP%20Report%20-%20version%206_mt%20revisions-2-5-13.pdf [March 2016].

Belzer, M.H. (2000). *Sweatshops on Wheels: Winners and Losers in Trucking Deregulation*. New York: Oxford University Press.

Belzer, M.H., Rodriguez, D.A., and Targa, F. (2006). Pay incentives and truck driver safety: A case study. *Industrial and Labor Relations Review*, *59*(2), 205-225.

Belzowski, B.M., Blower, D., Woodrooffe, J., and Green, P.E. (2007). *Tracking the Use of Onboard Safety Technologies across the Truck Fleet* (UMTRI 2009-22). Ann Arbor: University of Michigan Transportation Research Institute.

Berger, M., Varvarigou, Reilly, A., Czeisler, C.A., Maholtra, A., and Kales, S.N. (2012). Employer-mandated sleep apnea screening and diagnosis in commercial drivers. *Journal of Occupational and Environmental Medicine, 54*(8), 1017-1025.

Biglan, A., Ary, D., and Wagenaar, A.C. (2000). The value of interrupted time-series experiments for community intervention research. *Prevention Science*, *1*(1), 31-49.

Blanco, M., Hanowski, R.J., Olson, R.L., Morgan, J.F., Soccolich, S.A., Wu, S.-C., and Guo, F. (2011). *The Impact of Driving, Non-Driving Work, and Rest Breaks on Driving Performance in Commercial Motor Vehicle Operations* (FMCSA-RRR-11-017). Washington, DC: U.S. Department of Transportation, Federal Motor Carrier Safety Administration.

Blood, R.P., Yost, M.G., Camp, J.E., and Ching, R.P. (2015). Whole-body vibration exposure intervention among professional bus and truck drivers: A laboratory evaluation of seat-suspension designs. *Journal of Occupation and Environmental Hygiene, 12*(6), 351-362.

Bloom, H. (2012). Modern regression discontinuity analysis. *Journal of Research on Educational Effectiveness*, *5*(1), 43-82.

Blower, D., and Campbell, K.L. (2002). *The Large Truck Crash Causation Study* (UMTRI 2002-31). Washington, DC: U.S. Department of Transportation, Federal Highway Administration Office of Motor Carriers.

Bonauto, D.K., Lu, D., and Fan, Z.J. (2014). Obesity prevalence by occupation in Washington state, Behavioral Risk Factor Surveillance System. *Preventing Chronic Disease*, *11*, 130219.

Bonnet, M.H. (1991). The effect of varying prophylactic naps on performance, alertness and mood throughout a 52-hour continuous operation. *Sleep*, *14*(4), 307-315.

Bonnet, M.H. (2005). Acute sleep deprivation. In M.A. Kryger, T. Roth, and W.C. Dement (Eds.), *Principles and Practice of Sleep Medicine* (pp. 51-66). Philadelphia, PA: Elsevier Saunders.

Borbély, A.A. (1982). A two process model of sleep regulation. *Human Neurobiology, 1*(3), 195-204.

Braver, E.R., Preusser, C.W., Preusser, D.F., Baum, H.M., Beilock, R., and Ulmer, R. (1992). Long hours and fatigue: A survey of tractor-trailer drivers. *Journal of Public Health Policy*, *13*(3), 341-366.

Brooks, A., and Lack, L. (2006). A brief afternoon nap following nocturnal sleep restriction: Which nap duration is most recuperative? *Sleep*, *29*(6), 831-840.

Brown, T., Lee, J., Schwarz, C., Dary Fiorentino, D., and McDonald, A. (2014). *Assessing the Feasibility of Vehicle-Based Sensors to Detect Drowsy Driving*. Report No. DOT HS 811 886. Washington, DC: National Highway Traffic Safety Administration.

Burke, T.M., Scheer, F., Ronda, J.M., Czeisler, C.A., and Wright, K.P. (2015). Sleep inertia, sleep homeostatic and circadian influences on higher-order cognitive functions. *Journal of Sleep Research, 24*(4), 364-371.

Burks, S.V., Belzer, M.H., Quon, K., Pratt, S.G., and Shackelford, S. (2010). *Trucking 101: An Industry Primer* (Transportation Research Circular E-C146). Washington, DC: Transportation Research Board.

Buxton, O.M., Lee, C.W., L'Hermite-Baleriaux, M., Turek. F.W., and Van Cauter, E. (2003). Exercise elicits phase shifts and acute alterations of melatonin that vary with circadian phase. *American Journal of Physiology*, *284*(3), R714-R724.

Buxton, O.M., Pavlova, M., Reid, E.W., Wang, W., Simonson, D.C., and Adler, G.K. (2010). Sleep restriction for 1 week reduces insulin sensitivity in healthy men. *Diabetes*, *59*(9), 2126-2133.

Buysee, D.J., Reynolds, C.F., Monk, T.H., Berman, S.R., and Kupfer, D.J. (1989). The Pittsburgh Sleep Quality Index: A new instrument for psychiatric research and practice. *Psychiatry Research*, *28*(2), 193-213.

Caldwell, J.A., Prazinko, B., and Caldwell, J.L. (2003). Body posture affects electroencephalographic activity and psychomotor vigilance task performance in sleep-deprived subjects. *Clinical Neurophysiology*, *114*(1), 2-31.

Cannon, B.R., and Sudweeks, J. (2011). *Geospatial Analysis of High-Crash Intersections and Rural Roads Using Naturalistics Driving Data: Final Report* (11-UT-013). Blacksburg, VA: Virginia Tech Transportation Institute, National Surface Transportation Safety Center for Excellence.

Carayon, P. (2006). Human factors of complex sociotechnical systems. *Applied Ergonomics*, *37*(4), 525-535.

Carson, J., Pezoldt, V., Koncz, N., and Obeng-Boampong, K. (2011). *Benefits of Public Roadside Safety Rest Areas in Texas: Technical Report*. Arlington: Texas Transportation Institute, Texas Department of Transportation.

Cassel, W., Ploch, T., Becker, C., Dugnus, D., Peter, J.H., and von Wichert, P. (1996). Risk of traffic accidents in patients with sleep-disordered breathing: Reduction with nasal CPAP. *European Respiratory Journal*, *9*(12), 2606-2611.

Centers for Disease Control and Prevention. (2015). *Insufficient Sleep Is a Public Health Problem*. Available: http://www.cdc.gov/features/dssleep [March 2016].

Chakraborty, B., and Murphy, S.A. (2013). Dynamic treatment regimes. *Annual Review of Statistics and Its Application*, *1*, 447-464.

Charlton, S.G., and Ashton, M.E. (1998). *Review of Fatigue Management Strategies in the Transport Industry*. Wellington, New Zealand: Land Transport Safety Authority.

Chua, E.C., Tan, W.Q., Yeo, S.C., Lau, P., Lee, I., Mien, I.H., Puvanendran, K., and Gooley, J.J. (2012). Heart rate variability can be used to estimate sleepiness-related decrements in psychomotor vigilance during total sleep deprivation. *Sleep, 35*(3), 325-334.

Cohen, D.A., Wang, W., Wyatt, J.K., Kronauer, R.E., Dijk, D.-J., Czeisler, C.A., and Klerman, E.B. (2010). Uncovering residual effects of chronic sleep loss on human performance. *Science Translational Medicine*, *2*(14), 1-9.

Cole, S.R., and Hernan, M.A. (2008). Constructing inverse probability weights for marginal structural models. *American Journal of Epidemiology*, *168*(6), 656-664.

Connor, J., Norton, R., Ameratunga, S., Robinson, E., Civil, I., Dunn, R., Bailey, J., and Jackson, R. (2002). Driver sleepiness and risk of serious injury to car occupants: Population based case control study. *British Medical Journal*, *324*(7346), 1125.

Cook, T.D. (1991). Clarifying the warrant for generalized causal inferences in quasi-experimentation. In M.W. McLaughlin and D. Phillips (Eds.), *Evaluation and Education: At Quarter Century* (pp. 115-144). Chicago: National Society for the Study of Education.

Costa, G. (1997). The problem: Shiftwork. *Chronobiology International*, *14*(2), 89-98.

Couper, F.J., Pemberton, M., Jarvis, A., Hughes, M., and Logan, B.K. (2002). Prevalence of drug use in commercial tractor-trailer drivers. *Journal of Forensic Sciences*, *47*(3), 562-567.

Crouch, D.J., Birky, M.M., Gust, S.W., Rollins, D.E., Walsh, J.M., Moulden, J.V., Quinlan, K.E., and Beckel, R.W. (1993). The prevalence of drugs and alcohol in fatally injured truck drivers. *Journal of Forensic Sciences*, *38*(6), 1342-1353.

Crum, M.R., Morrow, P.C., and Daecher, C.W. (2002). *Motor Carrier Scheduling Practices and Their Influence on Driver Fatigue* (FMCSA-RT-03-005). Washington, DC: U.S. Department of Transportation, Federal Motor Carrier Safety Administration.

Czeisler, C. (2015). Duration, timing and quality of sleep are each vital for health, performance and safety. *Sleep Health*, *1*(1), 5-8.

Darwent, D., Roach, G., and Dawson, D. (2012). How well do truck drivers sleep in cabin sleeper berths?, *Applied Ergonomics*, *43*(2), 442-446.

Datta, T.K., Savolainen, P.T., Gates, T.J., Kay, J.J., Nicita, N.B., Sahadev, P., and Finkelman, J. (2015). *Evaluation of Non-Freeway Rumble Strips—Phase II* (RC-1627). Detroit, MI: Wayne State University Transportation Research Group.

Dawson, D., Noy I., Harma, M., Akerstedt, T., and Belenky, G. (2011). Modelling fatigue and the use of fatigue models in work settings. *Accident Analysis and Prevention, 43*(2), 549-564.

Day, R., Gerhardstein, R., Lumley, A., Roth, T., and Rosenthal, L. (1999). The behavioral morbidity of obstructive sleep apnea. *Progression of Cardiovascular Disease*, *41*(5), 341-354.

Deaconson, T.F., O'Hair, D.P., Levy, M.F., Lee, M.B.F., Schueneman, A.L., and Condon, R.E. (1988). Sleep deprivation and resident performance. *Journal of the American Medical Association*, *260*(12), 1721-1727.

Dement, W., Carskadon, M., and Richardson, G. (1978). Excessive daytime sleepiness in the sleep apnea syndrome. In C. Guilleminault and W. Dement (Eds.), *The Sleep Apnea Syndromes* (pp. 23-46). New York: R. Liss.

Dijkman, M., Sachs, N., Levine, E., Mallis, M., Carlin, M.M., Gillen, K.A., Powell, J.W., Samuel, S., Mullington, J., Rosekind, M.R., and Dinges, D.F. (1997). Effects of reduced stimulation on neurobehavioral alertness depend on circadian phase during human sleep deprivation. *Sleep*, *26*, 265.

Dinges, D.F. (1986). Differential effects of prior wakefulness and circadian phase on nap sleep. *Electroencephalography and Clinical Neurophysiology, 64*(3), 224-227.

Dinges, D.F. (1990). Are you awake? Cognitive performance and reverie during the hypnopompic state. In R.R. Bootzin, J.F. Kahlstrom, and D.L. Schacter (Eds.), *Sleep and Cognition* (pp. 139-175). Washington, DC: American Psychology Society.

Dinges, D.F., and Mallis, M.M. (1998). Managing fatigue by drowsiness detection: Can technological promises be realized?. In L. Hartley (Ed.), *Managing Fatigue in Transportation* (Ch. 11) (pp. 209-229). Oxford, UK: Elsevier Science.

Dinges, D.F., and Powell, J.W. (1985). Microcomputer analyses of performance on a portable, simple visual RT task during sustained operations. *Behavior Research Methods, Instruments, & Computers*, *17*(6), 652-655.

Dinges, D.F., Orne, M.T., Whitehouse, W.G., and Orne, E.C. (1987). Temporal placement of a nap for alertness: Contributions of circadian phase and prior wakefulness. *Sleep, 10*(4), 313-329.

Dinges, D.F., Pack, F., Williams, K., Gillen, K.A., Powell, J.W., Ott, G.E., Aptowicz, C., and Pack, A.I. (1997). Cumulative sleepiness, mood disturbance, and psychomotor vigilance performance decrements during a week of sleep restricted to 4-5 hours per night. S*leep, 20*(4), 267-277.

Dinges, D.F., Mallis, M.M., Maislin, G., and Powell, J.W. (1998). *Evaluation of Techniques for Ocular Measurement as an Index of Fatigue and the Basis for Alertness Management* (DOT HS 808 762). Available: http://ntl.bts.gov/lib/21000/21900/21955/PB99150237.pdf [March 2016].

Dinges, D.F., Price, N.J., Maislin, G., Powell, J.W., Ecker, A.J., Mallis, M.M., and Szuba, M.P. (2002). *Prospective Laboratory Revalidation of Ocular-Based Drowsiness Detection Technologies and Countermeasures* (NHTSA Drowsy Driver Detection and Interface Project DTNH-22-00-D-07007). Washington, DC: National Highway Traffic Safety Administration.

Dinges, D.F., Rider, R.L., Dorrian, J., McGlinchey, E.L., Rogers, N.L., Cizman, Z., Goldenstein, S.K., Vogler, C., Venkataraman, S., and Metaxas, D.N. (2005a). Optical computer recognition of facial expressions associated with stress induced by performance demands. *Aviation, Space, and Environmental Medicine*, *76*(6), B172-B182.

Dinges, D.F., Maislin, G. Brewster, R.M., Krueger, G.P., and Carroll, R.J. (2005b). Pilot test of fatigue management technologies. *Transportation Research Record: Journal of the Transportation Research Board, 1922*, 175-182.

Dinges, D.F., Venkataraman, S., McGlinchey, E.L., and Metaxas, D.N. (2007). Monitoring of facial stress during space flight: Optical computer recognition combining discriminative and generative methods. *Acta Astronautica, 60*(4-7), 341-350.

Dingus, T.A., Klauer, S.G., Neale, V.L., Petersen, A., Lee, S.E., Sudweeks, J., Perez, M.A., Hankey, J., Ramsey, D., Gupta, S., Bucher, C., Doerzaph, Z.R., Jermeland, J., and Knipling, R.R. (2006). *The 100-Car Naturalistic Driving Study: Phase II—Results of the 100-Car Field Experiment* (DOT HS 810 593). Washington, DC: National Highway Traffic Safety Administration.

Doll, R., and Hill, A.B. (1950). Smoking and carcinoma of the lung: Preliminary report. *British Medical Journal, 221*, 739-748.

Dozza, M., Werneke, J., and Mackenzie, M. (2013). e-BikeSAFE: A naturalistic cycling study to understand how electrical bicycles change cycling behaviour and influence safety. *Journal of Trauma*, *12*, 193-207.

Driskel, J.E., and Mullen, B. (2005). The efficacy of naps as a fatigue countermeasure: A meta-analytic integration. *Human Factors*, *47*(2), 360-377.

Drummond, S.P., Brown, G.G., Gillin, J.C., Stricker, J.L., Wong, E.C., and Buxton, R.B. (2000). Altered brain response to verbal learning following sleep deprivation. *Nature, 403*(6770), 655-657.

Duncan, G.T., Elliot, M., and Salazar-Gonzalez, J.-J. (2011). *Statistical Confidentiality*. New York: Springer.

Durand, G., and Kales, S.N. (2009). Obstructive sleep apnea screening during commercial driver medical examinations: A survey of ACOEM members. *Journal of Occupational and Environmental Medicine*, *51*(10), 1220-1226.

Eastman, C.I., Boulos, Z., Terman, M., Campbell, S.S., Dijk, D., and Lewy, A.J. (1995). Light treatment for sleep disorders: Consensus report. VI: Shift work. *Journal of Biological Rhythms*, *10*(2), 157-164.

Engleman, H.M., Hirst, W.S., and Douglas, N.J. (1997). Under reporting of sleepiness and driving impairment in patients with sleep apnoea/hypopnoea syndrome. *Journal of Sleep Research*, *6*(4), 272-275.

Federal Motor Carrier Safety Administration. (2002). *Cardiovascular Advisory Panel Guidelines for the Medical Examination of Commercial Motor Vehicle Drivers* (FMCSA-MCP-02-002). Washington, DC: U.S. Department of Transportation, Federal Motor Carrier Safety Administration.

Federal Motor Carrier Safety Administration. (2006). *Report to Congress on the Large Truck Crash Causation Study*. Available: https://www.fmcsa.dot.gov/sites/fmcsa.dot.gov/files/docs/ltccs-2006.pdf [March 2016].

Federal Motor Carrier Safety Administration. (2008). *Use of Income Derived from the Commercial Driver's License Information System for Modernization* (MH-2008-059). Available: https://www.oig.dot.gov/sites/default/files/CDLIS_Final_03.pdf [March 2016].

Federal Motor Carrier Safety Administration. (2014). *Pocket Guide to Large Truck and Bus Statistics*. Available: https://www.fmcsa.dot.gov/sites/fmcsa.dot.gov/files/docs/FMCSA%20Pocket%20Guide%20to%20Large%20Truck%20and%20Bus%20Statistics%20-%202014%20-%20508C.pdf [March 2016].

Federal Motor Carrier Safety Administration Medical Review Board. (2008). *Summary for the January 28, 2008, Medical Review Board Public Meeting*. Available: https://www.fmcsa.dot.gov/summary-january-28-2008-medical-review-board-public-meeting [March 2016].

Federal Motor Carrier Safety Administration Medical Review Board. (2012). *February 6, 2012 MCSAC and MRB Task 11-05: Final Report on Obstructive Sleep Apnea (OSA)*. Available: https://www.fmcsa.dot.gov/february-6-2012-mcsac-and-mrb-task-11-05-final-report-obstructive-sleep-apnea-osa [March 2016].

Ferrie, J.E., Shipley, M.J., Cappuccio, F.P., Brunner, E., Miller, M.A., Kumari, M., and Marmot, M.G. (2007). A prospective study of change in sleep duration: Associations with mortality in the Whitehall II cohort. *Sleep*, *30*(12), 1659-1666.

Feuerstein, C., Naegele, B., Pepin, J.L., and Levy, P. (1997). Frontal lobe-related cognitive functions in patients with sleep apnea syndrome before and after treatment. *Acta Neurologica Belgica*, *97*(2), 96-107.Findley, L., Smith, C., Hooper, J., Dineen, M., and Suratt, P.M. (2000). Treatment with nasal CPAP decreases automobile accidents in patients with sleep apnea. *American Journal of Respiratory Critical Care Medicine*, *161*(3, Pt. 1), 857-859.

Findley, L., Smith, C., Hooper, J., Dineen, M., and Suratt, P.M. (2000). Treatment with nasal CPAP decreases automobile accidents in patients with sleep apnea. *American Journal of Respiratory and Critical Care Medicine, 161*(3), 857-859.

Ford, E.S., Cunningham, T.J., and Croft, J.B. (2015). Trends in self-reported sleep duration among US adults from 1985 to 2012. *Sleep*, *38*(5), 829-832.

Fourie, C., Holmes, A., Bourgeois-Bougrine, S., Hilditch, C., and Jackson, P. (2010a). *Fatigue Risk Management Systems: A Review of the Literature* (Road Safety Research Report No. 110). London, UK: Clockwork Research.

Fourie, C., Holmes, A., Hilditch, C., Bourgeois-Bougrine, S., and Jackson, P. (2010b). *Interviews with Operators, Regulators and Researchers with Experience of Implementing Fatigue Risk Management Systems* (Road Safety Research Report No. 120). London, UK: Clockwork Research.

Gail, M.H. (1996). Statistics in action. *Journal of the American Statistical Association*, *91*(433), 1-13.Gander, P.H. (2015). Evolving regulatory approaches for managing fatigue risk in transport operations. *Reviews of Human Factors and Ergonomics*, *10*(1), 253-271.

Gander, P.H. (2015). *Fatigue, Safety, and Health: New Challenges and New Solutions.* Massey University, University of New Zealand. Available: http://www.ppta.org.nz/membershipforms/doc_view/2001-fatigue-safety-and-health-new-challenges-and-new-solutions-philippa-gander [April 2016}.

Gander, P.H., Millar, M., Webster, C., and Merry, A. (2008). Sleep loss and performance of anaesthesia trainees and specialists. *Chronobiology International, 25*(6), 1077-1091.

Garbarino, S., Nobili, L., Beelke, M., De Carli, F., Balestra, V., and Ferrillo, F. (2001). Sleep related vehicle accidents on Italian highways. *Giornale Italiano di Medicina del Lavoro ed Ergonomia*, *23*(4), 430-434.

Garder, P., and Davies, M. (2006). Safety effect of continuous shoulder rumble strips on rural interstates in Maine. *Transportation Research Record: Journal of the Transportation Research Board*, *1853*, 7.

Gay, P., Weaver, T., Loube, D., and Iber, C. (2006). Evaluation of positive airway pressure treatment for sleep related breathing disorders in adults: A review by the Positive Airway Pressure Task Force of the Standards of Practice Committee of the American Academy of Sleep Medicine. *Sleep*, *29*(3), 381-401.

George, C.F., and Smiley, A. (1999). Sleep apnea and automobile crashes. *Sleep*, *22*(6), 790-795.

George, C.F., and Smiley, A. (2001). Reduction in motor vehicle collisions following treatment of sleep apnoea with nasal CPAP. *Thorax*, *56*(7), 508-512.

Gillberg, M. (1984). The effects of two alternative timings of a one-hour nap on early morning performance. *Biological Psychology*, *19*(1), 45-54.

Goel, A., and Vidal, T. (2014). Hours-of-service regulations in road freight transport: An optimization-based international assessment. *Transportation Science*, *48*(3), 391-412.

Goel, N., Rao, H., Durmer, J.S., and Dinges, D.F. (2009). Neurocognitive consequences of sleep deprivation. *Seminars in Neurology, 29*(4), 320-339.

Goel, N., Banks, S., Mignot, E., and Dinges, D.F. (2010). DQB1*0602 predicts interindividual differences in physiologic sleep, sleepiness, and fatigue. *Neurology, 75*, 1509-1519.

Greenland, S. (1989). Modeling and variable selection in epidemiologic analysis. *American Journal of Public Health*, *79*(3), 340-349.

Guo, F., and Fang, Y. (2013). Individual driver risk assessment using naturalistic driving data. *Accident Analysis and Prevention*, *61*, 3-9.

Guo, F., Klauer, S.G., Hankey, J.M., and Dingus, T.A. (2010). Near crashes as crash surrogates for naturalistic driving studies. *Transportation Research Record: Journal of the Transportation Research Board*, *2147*, 66-74.

Haddon, W., Jr. (1972). A logical framework for categorizing highway safety phenomena and activity. *Journal of Trauma, 12*(3), 193-207.

Hanowski, R.J. (2013). Behavioural adaptation and unintended consequences: Examples from commercial vehicle operations. In C. Rudin-Brown and S. Jamson (Eds.), *Behavioural Adaptation and Road Safety: Theory, Evidence and Action* (pp. 323-336). Boca Raton, FL: Taylor & Francis Group (CRC Press).

Hanowski, R.J., Wierwille, W.W., Garness, S.A., and Dingus, T.A. (2000). *Impact of Local/Short Haul Operations on Driver Fatigue: Final Project Report* (DOT-MC-00-203). Washington, DC: U.S. Department of Transportation, Federal Motor Carrier Safety Administration.

Hanowski, R.J., Wierwille W., and Dingus T. (2003). An on-road study to investigate fatigue in local/short haul trucking. *Accident Analysis and Prevention, 35*(2), 153-160.

Hanowski, R.J., Hickman, J.S., Fumero, M.C., Olson, R.L., and Dingus, T.A. (2007). The sleep of commercial motor vehicle drivers under the 2003 hours-of-service regulations. *Accident Analysis and Prevention*, *39*(6), 1140-1145.

Hanowski, R.J., Olson, R.L., Bocanegra, J., and Hickman, J.S. (2008). *Analysis of Risk as a Function of Driving-Hour: Assessment of Driving-Hours 1 through 11* (FMCSA RRR-08-002). Washington, DC: U.S. Department of Transportation, Federal Motor Carrier Safety Administration.

Hanowski, R.J., Hickman, J.S., Olson, R.L., and Bocanegra, J. (2009). Evaluating the 2003 revised hours-of-service regulations for truck drivers: The impact of time-on-task on critical incident risk. *Accident Analysis and Prevention*, *41*(2), 268-275.

Hartenbaum, N.P. (2015). Certified medical examiners and screening for obstructive sleep apnea. *Journal of Occupational and Environmental Medicine*, *57*(3), e19-e22.

Hartenbaum, N.P., Collop, N., Rosen, I.M., Phillips, B., George, C.F., Rowley, J.A., Freedman, N., Weaver, T.E., Gurubhagavatula, I., Strohl, K., Leaman, H.M., Moffitt, G.L., and Rosekind, M.R. (2006). Sleep apnea and commercial motor vehicle operators: Statement from the Joint Task Force of the American College of Chest Physicians, American College of Occupational and Environmental Medicine, and the National Sleep Foundation. *Journal of Occupational and Environmental Medicine*, *48*(Suppl. 9), S4-S37.

Hartley, L., Horberry, T., Mabbott, N., and Krurger. G.P. (2000). *Review of Fatigue Detection and Prediction Technologies*. Melbourne, Australia: Murdoch University Institute for Research in Safety and Transport, Australian National Road Transport Commission.

Hauer, E. (1995). On exposure and accident rate. *Traffic Engineering and Control, 36*(3), 134-138.

Hayashi, M., Masuda, A., and Hori, T. (2003). The alerting effects of caffeine, bright light and face washing after a short daytime nap. *Clinical Neurophysiology*, *114*(12), 2268-2278.

He, Y., Jones, C.R., Fujiki, N., Xu, Y., Guo, B., Holder, J.L., Jr., Rossner, M.J., et al. (2009). The transcriptional repressor DEC2 regulates sleep length in mammals. *Science, 325*(5942), 866-870.

Hebert, M. (2014). *Feature Extraction from Videos.* Presentation at the Second Meeting of the Panel on Research Methodologies and Statistical Approaches to Understanding Driver Fatigue Factors in Motor Carrier Safety and Driver Health, May 28-29, National Research Council, Washington, DC.

Hedlund, J., and Blower, D. (2006). *Large Truck Crash Causation Study (LTCCS) Analysis Series: Using LTCCS Data for Statistical Analysis of Crash Risk* (FMCSA-RI-05-037). Washington, DC: Federal Motor Carrier Safety Administration.

Hege, A., Perko, M., Johnson, A., Yu, C.H., Sönmez, S., and Apostolopoulos, Y. (2015). Surveying the impact of work hours and schedules on commercial motor vehicle driver sleep. *Safety Health Work*, *6*(2), 104-113.

Hernán, M.A., and Robins, J.M. (2006). Instruments for causal inference: An epidemiologist's dream? *Epidemiology*, *17*(4), 360-372.

Hernán, M.A., and Robins, J.M. (2008). Observational studies analyzed like randomized experiments: Best of both worlds. *Epidemiology, 19*(6), 789-792.

Heslegrave, R.J., and Angus, R.G. (1985). The effects of task duration and work-session location on performance degradation induced by sleep loss and sustained cognitive work. *Behavioral Research Methods, Instruments and Computers*, *17*(6), 592-603.

Hickman, J.S., and Hanowski, R.J. (2011). Use of a video monitoring approach to reduce at risk driving behaviors in commercial vehicle operations. *Transportation Research Part F: Traffic Psychology and Behavior, 14*(3), 188-198.

Hill, A.B. (1965). The environment and disease: Association or causation? *Proceedings of the Royal Society of Medicine*, *58*(5), 295-300.

Hill, J.L. (2011). Bayesian nonparametric modeling for causal inference. *Journal of Computational and Graphical Statistics*, *20*(1), 217-240.

Hirano, K., Guido, I.W., Ridder, G., and Rubin, D.B. (2001). Combining panel data sets with attrition and refreshment samples. *Econometrica*, *69*(6), 1645-1659.

Hoenig, J.M., and Heisey, D.M. (2001). The abuse of power: The pervasive fallacy of power calculations in data analysis. *The American Statistician*, *55*(1), 19-24.

Holland, P.W. (1986). Statistics and causal inference. *Journal of the American Statistical Association*, *81*(396), 945-960.

Holland, P.W. (1988). Causal mechanism or causal effect? Which is best for Statistical Science? *Statistical Science, 3*(2), 186-188.

Horne, J.A., and Foster, S.C. (1995). Can exercise overcome sleepiness? *Sleep Research*, *24A*, 437.

Horne, J.A., and Reyner, L.A. (1995). Sleep related vehicle accidents. *British Medical Journal*, *310*(6979), 565-567.

Hortsmann, S., Hess, C.W., Bassetti, C., Gugger, M., and Mathis, J. (2000). Sleepiness-related accidents in sleep apnea patients. *Sleep*, *23*(3), 383-389.

Hsu, J.Y., and Small, D.S. (2013). Calibrating sensitivity analyses to observed covariates in observational studies. *Biometrics*, *69*(4), 803-811.

Huang, W., Ramsey, K.M., Marcheva, B., and Bass, J. (2011). Circadian rhythms, sleep, and metabolism. *Journal of Clinical Investigation*, *121*, 2133-2141.

Hursh, S.R., Redmond, D.P., Johnson, M.L., Thorne, D.R., Belenky, G., Balkin, T.J., Storm, W.F., Miller, J.C., and Eddy, D.R. (2004). Fatigue models for applied research in warfighting. *Aviation, Space, and Environmental Medicine, 75*(3), 44-53.

Institute of Medicine. (2006). *Sleep Disorders and Sleep Deprivation: An Unmet Public Health Problem.* H.R. Colten and B.M. Altevogt (Eds.). Committee on Sleep Medicine and Research. Board on Health Sciences Policy. Washington, DC: The National Academies Press.

Jackson, J., Albert, P.S., Zhang, Z., and Simons-Morton, B. (2013). Ordinal latent variable models and their application in the study of newly licensed teenage drivers. *Journal of the Royal Statistical Society: Series C*, *62*(3), 435-450.

Jackson, J.C., Albert, P.S., and Zhang, Z. (2015). A two-state mixed hidden Markov model for risky teenage driving behavior. *Annals of Applied Statistics*, *9*, 849-865.

Jewett, M.E., Wyatt, J.K., Ritz-De Cecco, A., Khalsa, S.B., Dijk, D.J., and Czeisler, C.A. (1999). Time course of sleep inertia dissipation in human performance and alertness. *Journal of Sleep Research*, *8*(1), 1-8.

Joffe, M.M., and Greene, T. (2009). Related causal frameworks for surrogate outcomes. *Biometrics*, *65*(2), 530-538.

Jonasson, J.K., and H. Rootzén. (2014). Internal validation of near-crashes in naturalistic driving studies: A continuous and multivariate approach. *Accident Analysis and Prevention*, *62*, 102-109.

Jones, I., and Stein, H. (1989). Defective equipment and tractor-trailer crash involvement. *Accident Analysis and Prevention*, *21*(5), 469-481.

Jordan, A., McSharry, D., and Malhotra, A. (2014). Adult obstructive sleep apnea. *The Lancet, 22*, 736-747.

Jovanis, P.P., Wu, K.F., and Chen, C. (2011). *Hours of Service and Driver Fatigue: Driver Characteristic Research* (FMCSA-RRR-11-018). Washington, DC: Federal Motor Carrier Safety Administration.

Kaneko, Y., Hajek, V.E., Zivanovic, V., Raboud, J., and Bradley, T.D. (2003). Relationship of sleep apnea to functional capacity and length of hospitalization following stroke. *Sleep*, *26*(3), 293-297.

Khan, M., Abdel-Rahim, A., and Williams, C.J. (2015). Potential crash reduction benefits of shoulder rumble strips in two-lane rural highways. *Accident Analysis and Prevention*, *75*, 35-42.

Kim, H.C., Young, T., Matthews, C. G., Weber, S.M., Woodard, A.R., and Palta, M. (1997). Sleep-disordered breathing and neuropsychological deficits: A population-based study. *American Journal of Respiratory and Critical Care Medicine*, *156*(6), 1813-1819.

Kim, S., Chen, Z., Simons-Morton, B.G., and Albert, P.S. (2013). Bayesian hierarchical Poisson regression models: An application to a driving study with kinematic events. *Journal of the American Statistical Association*, *108*(502), 494-503.

Knipling, R.R. (2015). *Review of Commercial Driver Fatigue Research Methodologies*. Commissioned paper for the National Academies of Sciences, Engineering, and Medicine Panel on Research Methodologies and Statistical Approaches to Understanding Driver Fatigue Factors in Motor Carrier Safety and Driver Health, Washington, DC. Available: http://nas.edu/DriverFatigue_Knipling_Paper [March 2016].

Knipling, R.R., and Wang, J.-S. (1994). *Crashes and Fatalities Related to Driver Drowsiness/Fatigue*. Washington, DC: National Highway Traffic Safety Administration

Knipling, R.R., and Wang, J.-S. (1995). *Revised Estimates of the U.S. Drowsy Driver Crash Problem Size Based on General Estimates System Case Reviews*. 39th Annual Proceedings, Association for the Advancement of Automotive Medicine, October, Chicago, IL.

Kribbs, N.B., Pack, A.I., Kline, L.R., Smith, P.L., Schwartz, A.R., Schubert, N.M., Redline, S., Henry, J.N, Getsy, J.E., and Dinges, D.F. (1993). Objective measurement of patterns of nasal CPAP use by patients with obstructive sleep apnea. *American Review of Respiratory Disease*, *147*(4), 887-895.

Krueger, G.P. (2008). Health, wellness, fitness and commercial driver safety: A review of the issues. *Journal of the Washington Academy of Sciences*, *94*(3), 31-59.

Krueger, G.P. (2010a). Psychoactive medications, stimulants, hypnotics, and nutritional aids: Effects on driving alertness and performance. *Journal of the Washington Academy of Sciences*, *96*(4), 51-85.

Krueger, G.P. (2010b). *Overview: Research on the Health and Wellness of Commercial Truck and Bus Drivers*. Summary of an International Conference, November 8-10, Baltimore, MD.

Krueger, G.P. (2012). *Research on the Health and Wellness of Commercial Truck and Bus Drivers: Summary of an International Conference*. Washington DC: Transportation Research Board. Available: http://onlinepubs.trb.org/onlinepubs/conf/cpw5.pdf [April 2016].

Krueger, G.P., and Brewster, R.M. (2002). *Gettin'-in-Gear: Wellness, Health, and Fitness Program for Commercial Drivers: Instructors Manual*. Alexandria, VA: American Trucking Associations' American Transportation Research Institute and Federal Motor Carrier Safety Administration.

Krueger, G.P., and Van Hemel, S.B. (1997). *Behavioral Task Analysis of Truck Drivers Loading and Unloading Trucks: Relationship to Driver Fatigue, Alertness, and Driving Safety*. Final Report.

Krueger, G.P., Brewster, R.M., Dick, V.R., Inderbitzen, R., and Staplin, L. (2007). *Health and Wellness Programs for Commercial Drivers* (Commercial Truck and Bus Safety Synthesis 15). Washington, DC: Transportation Research Board.

Krueger, G.P., Leaman, H.M., and Bergoffen, G. (2011). *Effects of Psychoactive Chemicals on Commercial Driver Health and Performance: Stimulants, Hypnotics, Nutritional, and other Supplements* (Commercial Truck and Bus Safety Synthesis 19). Washington, DC: Transportation Research Board.

Kucjarczyk, E.R., Morgan, K., and Hall, A.P. (2012). The occupational impact of sleep quality and insomnia symptoms. *Sleep Medicine Reviews, 16*(6), 547-559.

Lavie, P. (1986). Ultrashort sleep-waking schedule. III: "Gates" and "forbidden zones" for sleep. *Electroencephalography and Clinical Neurophysiology*, *63*(5), 414-425.

LeDuc, P.A., Caldwell, J.A., and Ruyak, P.S. (2000). *The Effects of Exercise versus Napping on Alertness and Mood in Sleep-Deprived Aviators* (Technical Report 2000-2012). Fort Rucker, AL: U.S. Aeromedical Research Laboratory.

Léger, D., and Bayon, V. (2010). Societal costs of insomnia. *Sleep Medicine Reviews, 14*(6), 379-389.

Lerman, S.E., Eskin, E., Flower, D.J., George, E.C., Gerson, B., Hartenbaum, N., Hursh, S.R., and Moore-Ede, M. (2012). ACOEM guidance statement: Fatigue risk management in the workplace. *Journal of Occupational and Environmental Medicine*, *54*(2), 231-258.

Lim, J., and Dinges, D.F. (2008). Sleep deprivation and vigilant attention. *Annals of the New York Academy of Sciences*, *1129*, 305-322.

Lim, J., and Dinges, D.F. (2010). A meta-analysis of the impact of short-term sleep deprivation on cognitive variables. *Psychological Bulletin*, *136*(3), 375-389.

Lim, J., Wu, W-C., Wang, J., Detre, J.A., Dinges, D.F., and Rao, H. (2010). Imaging brain fatigue from sustained mental workload: An ASL perfusion study of the time-on-task effect. *NeuroImage*, *49*(4), 3426-3435.

Liu, W., Kuramoto, S.J., and Stuart, E.A. (2013). An introduction to sensitivity analysis for unobserved confounding in nonexperimental prevention research. *Prevention Science*, *14*(6), 570-580.

Lund, A.K., Preusser, D.F., Blomberg, R.D, and Williams, A.F. (1988). Drug use by tractor-trailer drivers. *Journal of Forensic Sciences*, *33*(3), 648-661.

Mackie, R.R., and Miller, J.C. (1978). *Effects of Hours of Service Regularity of Schedules, and Cargo Loading on Truck and Bus Driver Fatigue*. Washington, DC: National Highway Traffic Safety Administration and Bureau of Motor Carrier Safety, U.S. Department of Transportation.

Maclure, M., and Mittleman, M.A. (2000). Should we use a case-crossover design? *Annual Review of Public Health*, *21*, 193-221.

Mallis, M.M., and Dinges, D.F. (2005). Monitoring alertness by eyelid closure. In N. Stanton, A. Hedge, K. Brookhuis, E. Salas, and H. Hendrick (Eds.), *The Handbook of Human Factors and Ergonomics Methods* (pp. 25.1-25.6). New York: CRC Press.

Mallis, M.M., and James, F.O. (2012). The role of alertness monitoring in sustaining cognition during sleep loss. In N.J. Wesensten and T.J. Balkin (Eds.), *Sleep Deprivation, Stimulant Medications, and Cognition* (Ch. 15) (pp. 209-222). New York: Cambridge University Press.

Mallis, M.M., Maislin, G., Konowal, N., Byrne, V., Bierman, D., Davis, R., Grace, R., and Dinges, D.F. (1998). *Biobehavioral Responses to Drowsy Driving Alarms and Alerting Stimuli*. Washington, DC: U.S. Department of Transportation.

Mallis, M.M., Mejdal, S., Nguyen, T.T., and Dinges, D.F. (2004). Summary of the key features of seven biomathematical models of human fatigue and performance. *Aviation, Space, and Environmental Medicine, 75*(Suppl. 3), A4-A14.

Mallis, M.M., Banks, S., and Dinges, D.F. (2007). Sleep and circadian control of neurobehavioral functions. In R. Parasuraman and M. Rizzo (Eds.), *Neuroergonomics: The Brain at Work* (pp. 207-220). Oxford, UK: Oxford University Press.

Mannering, F., and Bhat, C. (2014). Analytic methods in accident research: Methodological frontier and future directions. *Analytic Methods in Accident Research*, 1, 1-22.

Matsumoto, Y., Mishima, K., Satoh, K., Shimizu, T., and Hishikawa, Y. (2002). Physical activity increases the dissociation between subjective sleepiness and objective performance levels during extended wakefulness in human. *Neuroscience Letters*, *326*(2), 133-136.

McArthur, A., Kay, J., and Savolainen, P. (2013). *The Effects of Public Rest Areas on Fatigue-Related Crashes*. Proceedings of the 92nd Annual Meeting of the Transportation Research Board, January, Washington, DC.

McCartt, A.T., Hellinga, L.A., and Solomon, M.G. (2008). Work schedules of long-distance truck drivers before and after 2004 hours-of-service rule change. *Traffic in Prevention*, *9*(3), 201-210.

McCauley, P., Kalachev, L.V., Smith, A.D., Belenky, G., Dinges, D.F., and Van Dongen, H.P. (2009). A new mathematical model for the homeostatic effects of sleep loss on neurobehavioral performance. *Journal of Theoretical Biology, 256*(2), 227-239.

McCauley, P., Kalachev, L.V., Mollicone, D.J., Banks, S., Dinges, D.F., and Van Dongen, H.P. (2013). Dynamic circadian modulation in a mathematical model for the effects of sleep and sleep loss on waking neurobehavioral performance. *Sleep*, 36(12), 1989-1999.

McClafferty, J.A., Hankey, J.M., and Perez, M.A. (2015). *Description of the SHR2 Naturalistic Database and the Crash, Near-Crash, and Baseline Data Sets Task Report.* Virginia Tech Transportation Institute, Blacksburg, VA. Prepared for The Strategic Highway Research Program 2 Transportation Research Board of The National Academies. Available: http://vtechworks.lib.vt.edu/bitstream/handle/10919/70850/SHRP_2_CrashNearCrashBaselineReport_4-25-16.pdf?sequence=1&isAllowed=y [April 2016].

McCormick, F., Kadzielski, J., Evans, B.T., Landrigan, C.P., Herndon, J., and Rubash, H. (2013). Fatigue optimization scheduling in graduate medical education: Reducing fatigue and improving patient safety. *Journal of Graduate Medical Education*, *5*(1), 107-111.

McDonald, A.D., Lee, J.D., Aksan, N.S., Dawson, J.D., Tippin, J., and Rizzo, M. (2013). The language of driving: Advantages and applications of symbolic data reduction for analysis of naturalistic driving data. *Transportation Research Record*, *2392*, 22-30.

McKnight-Eily, L.R., Liu, Y., Perry, G.S., Presley-Cantrell, L.R., Strine, T.W., Lu, H., and Croft, J.B. (2009). Perceived insufficient rest or sleep among adults—United States, 2009. *Journal of the American Medical Association*, *302*(23), 2532.

McKnight-Eily, L.R., Liu, Y., Croft, J.B., Perry, G.S., and Strine, T. (2011). *Unhealthy Sleep-Related Behaviors—12 States, 2009*. Atlanta, GA: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention.

Meyer, B.D. (1995). Natural and quasi-experiments in economics. *Journal of Business & Economic Statistics, 13*(2), 151-161.

Mitler, M.M., Carskadon, M.A., Czeisler, C.A., Dement, W.C., Dinges, D.F., and Graeber, R.C. (1988). Catastrophes, sleep, and public policy: Consensus report. *Sleep*, *11*(1), 100-109.

Mittleman, M.A., Maclure, M., Sherwood, J.B., Mulry, R.P., Tofler, G.H., Jacobs, S.C., Friedman, R., Benson, H., and Muller, J.E. (1995). Triggering of acute myocardial infarction onset by episodes of anger. Determinants of Myocardial Infarction Onset Study Investigators. *Circulation*, *92*(7), 1720-1725.

Mollicone, D.J., Van Dongen, H.P.A., and Dinges, D.F. (2007). Optimizing sleep/wake schedules in space: Sleep during chronic nocturnal sleep restriction with and without diurnal naps. *Acta Astronautica*, *60*, 354-361.

Mollicone, D.J., Van Dongen, H.P.A., Rogers, N.L., and Dinges, D.F. (2008). Response surface mapping of neurobehavioral performance: Testing the feasibility of split sleep schedules for space operations. *Acta Astronautica*, *63*(7-10), 833-840.

Mollicone, D.J., Van Dongen, H.P.A., Rogers, N.L., Banks, S., and Dinges, D.F. (2010). Time of day modulates the effects of chronic sleep restriction on neurobehavioral performance. *Aviation, Space and Environmental Medicine*, *81*(8), 735-744.

Monaco, K., and Williams E. (2000). Assessing the determinants of safety in the trucking industry. *Journal of Transportation and Statistics*, *3*(1), 69-79.

Moscovitch, A., Reimer, M., Heslegrave, R., Boivin, D., Hirshkowitz, M., Rhodes, W., and Kealey, M. (2006). *Development of a North-American Fatigue Management Program for Commercial Motor Carriers: Phase II (Pilot Study)*. Ottawa, ON: Canadian Sleep Institute.

Moses, L.N., and Savage, I. (1996). Identifying dangerous trucking firms. *Risk Analysis*, *16*(3), 351-358.

Mulgrew, A.T., Nasvadi, G., Butt, A., Cheema, R., Fox, N., Fleetham, JA, Ryan, CF., Cooper, P., and Ayas, N.T. (2008). Risk and severity of motor vehicle crashes in patients with obstructive sleep apnoea/hypopnoea. *Thorax*, *63*(6), 536-541.

Najm, W., Koopmann, J., Smith, J.D., and Brewer, J. (2010). *Frequency of Target Crashes for IntelliDrive Safety Systems* (DOT HS 811 381). Washington, DC: National Highway Traffic Safety Administration.

National Center for Health Statistics. (2015). *Health, United States, 2014, with Special Feature on Adults Aged 55-64.* Hyattsville, MD: National Center for Health Statistics.

National Highway Traffic Safety Administration and Federal Motor Carrier Safety Administration. (2006a). *Large Truck Crash Causation Study Analytical User's Manual*. Washington, DC: National Highway Traffic Safety Administration, Federal Motor Carrier Safety Administration.

National Highway Traffic Safety Administration and Federal Motor Carrier Safety Administration. (2006b). *Large Truck Crash Causation Study Codebook*. Available: https://s3.amazonaws.com/cmisst.media/LTCCS_Codebook_060106_v6.pdf [March 2016].

National Highway Traffic Safety Administration and Federal Motor Carrier Safety Administration. (2012). *Large Truck Crash Causation Survey Case Viewer*. Search criteria. Available: http://www-nass.nhtsa.dot.gov/nass/ltccs/SearchForm.aspx [March 2016].

National Registry of Certified Medical Examiners. (2012). *Medical Examiner Sample Training Handbook*. Available: http://nrcme.fmcsa.dot.gov/documents/ME_Training_Document.pdf [March 2016].

National Sleep Foundation. (2012). *2012 Sleep in America Poll: Transportation Workers' Sleep*. Available: https://sleepfoundation.org/sleep-polls-data/sleep-in-america-poll/2012-transportation-workers-and-sleep [March 2016].

National Transportation Safety Board. (1990). *Fatigue, Alcohol, Other Drugs, and Medical Factors in Fatal-to-the-Driver Heavy Truck Crashes, Safety Study* (NTSB/SS-90/01). Washington, DC: National Transportation Safety Board.

National Transportation Safety Board. (2009). *NTSB Safety Recommendation H-09-015 and 016*. Available: http://www.ntsb.gov/safety/safety-recs/RecLetters/H09_15_16.pdf [March 2016].

Nehlig, A. (1999). Are we dependent upon coffee and caffeine? A review on human and animal data. *Neuroscience and Biobehavioral Reviews*, *23*(4), 563-576.

Neri, D.F., Wiegmann, D., Stanny, R.R., Shappell, S.A., McCardie, A., and McKay, D.L. (1995). The effects of tyrosine on cognitive performance during extended wakefulness. *Aviation, Space, and Environmental Medicine*, *66*(4), 313-319.

Neri, D.F., Oyung, R.L., Colletti, L.M., Mallis, M.M., Tam, P.Y., and Dinges, D.F. (2002). Controlled breaks as a fatigue countermeasure on the flight deck. *Aviation, Space, and Environmental Medicine*, *73*(7), 654-664.

Newhouse, J.P., and McClellan, M. (1998). Econometrics in outcomes research: The use of instrumental variables. *Annual Review of Public Health*, *19*, 17-34.

Nowak, M., Kornhuber, J., and Meyrer, R. (2006). Daytime impairment and neurodegeneration in OSAS. *Sleep*, *29*(12), 1521-1530.

O'Neill, T.R., Krueger, G.P., Van Hemel, S.B., McGowan, A.L, and Rogers, W.C. (1999). Effects of cargo loading and unloading on truck driver alertness. *Transportation Research Record*, *1686*(99-0789), 42-48.

Ong, J.L., Asplund, C.L., Chia, T.T., and Chee, M.W. (2013). Now you hear me, now you don't: Eyelid closures as an indicator of auditory task disengagement. *Sleep*, *36*(12), 1867-1874.

OOIDA Foundation. (2015). *Owner-Operator and Professional Employee Driver Facts*. Available: *Owner-Operator and Professional Employee Driver Facts.* Available: http://www.ooida.com/OOIDA%20Foundation/RecentResearch/OOfacts.asp [March 2016].

Pack, A.I., Pack, A.M., Rodgman, E., Cucchiara, A., Dinges, D.F., and Schwab, C.W. (1995). Characteristics of crashes attributed to the driver having fallen asleep. *Accident Analysis and Prevention*, *27*(6), 769-775.

Pack, A.I., Dinges, D.F., and Maislin, G. (2002). *A Study of Prevalence of Sleep Apnea among Commercial Truck Drivers* (DOT-RT-02-030). Washington, DC: U.S. Department of Transportation, Federal Motor Carrier Safety Administration.

Pack, A.I., Maislin, G., Staley, B., Pack, F.M., Rogers, W., George, C.F.P., and Dinges, D.F. (2006). Impaired performance in commercial drivers: Role of sleep apnea and sleep duration. *American Journal of Respiratory and Critical Care Medicine*, *174*(4), 446-454.

Patel, R.B., Council, F.M., and Griffith, M.S. (2007). *Estimating Safety Benefits of Shoulder Rumble Strips on Two Lane Rural Highways in Minnesota: An Empirical Bayes Observational Before-After Study*. TRB 86th Annual Meeting: Compendium of Papers CD-ROM, Washington, DC.

Pearl, J., and Bareinboim, E. (2011). External validity: From do-calculus to transportability across populations. *Statistical Science*, *29*(4), 579-595.

Pellegrino, R., Kayakli, I.H., Goel, N., Cardinale, C.J., Dinges, D.F., Kuna, S.T., Maislin, et al. (2014). A novel BHLHE41 variant is associated with short sleep and resistance to sleep deprivation in humans. *Sleep, 37*(8), 1327-1336.

Punjabi, N. M., Bandeen-Roche, K., and Young, T. (2003). Predictors of objective sleep tendency in the general population. *Sleep*, *26*(6), 678-683.

Putcha, D., Blower, D., and Campbell, K.L. (2002). Bus accidents in the United States, 1995-1999. In A.G. Zacharia (Ed.), *International Truck and Bus Safety Research and Policy Symposium* (pp. 261-272). Knoxville, TN: University of Tennessee Center for Transportation Research and National Safety Council.

Quan, S.F., and Barger, L.K. (2015). Brief review: Sleep health and safety for transportation workers. *Southwest Journal of Pulmonary and Critical Care, 10*(3), 130-139.

Reason, J. (1990). *Human Error*. New York: Cambridge University Press.

Rizzo, M. (2011). Impaired driving from medical conditions: A 70-year-old-man trying to decide if he should continue driving. *Journal of the American Medical Association*, *305*(10), 1018-1026.

Robins, J.M., Hernan, M.A., and Brumback, B. (2000). Marginal structural models and causal inference in epidemiology. *Epidemiology*, *11*(5), 550-560.

Rodríguez, D.A., Targa, F., and Belzer, M.H. (2006). Pay incentives and truck driver safety: A case study. *Industrial and Labor Relations Review*, *59*(2), article 2.

Rogers, N.L., Dorrian, J., and Dinges, D.F. (2003). Sleep, waking and neurobehavioural performance. *Frontiers in Bioscience*, *8*, S1056-S1067.

Rosekind, M.R., Graeber, R.C., Dinges, D.F., Connell, L.J., Rountree, M., Spinweber, C.L., and Gillen, K.A. (1994). *Crew Factors in Flight Operations. IX: Effects of Cockpit Rest on Crew Performance and Alertness in Long-Haul Operations* (Report No. DOT/FAA/92/24). Moffett Field, CA: NASA Ames Research Center.

Rosenbaum, P.R. (1987). The role of a second control group in an observational study (with discussion). *Statistical Science*, *2*(3), 292-316.

Rosenbaum, P.R. (2002). *Observational Studies* (2nd Edition). New York: Springer-Verlag.

Rosenbaum, P.R. (2004). Design sensitivity in observational studies. *Biometrika*, *91*(1), 153-164.

Rosenbaum, P.R. (2005). Sensitivity analysis in observational studies. In B.S. Everitt and D.C. Howell (Eds.), *Encyclopedia of Statistics in Behavioral Science* (Vol. 4) (pp. 1809-1814). Chichester, UK: John Wiley.

Rosenbaum, P.R. (2009). Evidence factors in observational studies. *Biometrika*, *97*(2), 333-345.

Rosenbaum, P.R., and Rubin, D.B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, *70*(1), 41-55.

Rubin, D.B. (1997). Estimating causal effects from large data sets using propensity scores. *Annals of Internal Medicine*, *127*(8, Pt. 2), 757-763.

Ruehland, W.R., Rochford, P.D., O'Donoghue, F.J., Pierce, R.J., Singh, P., and Thornton, A.T. (2009). The new AASM criteria for scoring hypopneas impact on the apnea/hypopnea index. *Sleep*, *32*(2), 150-157.

Runyan, C. (1998). Using the Haddon matrix: Introducing the third dimension. *Injury Prevention*, *4*(4), 302-307.

Saltzman, G.M., and Belzer, M.H. (2007). *Truck Driver Occupational Safety and Health: 2003 Conference Report and Selective Literature Review* (2007-120). Washington, DC: U.S. Department of Health and Human Services, Public Health Service, Centers for Disease Control and Prevention, National Institute for Occupational Safety and Health, Department of Health and Human Services.

Sayed, T., P. deLeur, and J. Pump. (2010). *Impact of Rumble Strips on Collision Reduction on BC Highways: A Comprehensive Before and After Safety Study*. TRB 89th Annual Meeting Compendium of Papers CD-ROM, Washington, DC.

Schutte-Rodin, S., Broch, L., Buysee, D., Dorsey, C., and Satela, M. (2008). Clinical guideline for the evaluation and management of chronic insomnia in adults. *Journal of Clinical Sleep Medicine, 4*(5), 487-504.

Schweitzer, P.K., Muehlback, M.J., and Walsh, J.K. (1992). Countermeasures for night work performance deficits: The effect of napping or caffeine on continuous performance at night. *Work and Stress*, *6*(4), 355-365.

Scott, L.D., Hwang, W.T., Rogers, A.E., Nysse, T., Dean, G.E., and Dinges, D.F. (2007). The relationship between nurse work schedules, sleep duration, and drowsy driving. *Sleep*, *30*(12),1801-1807.

Shadish, W.R., Cook, T.D., and Campbell, D.T. (2002). *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Boston, MA: Houghton Mifflin.

Shaikh, S., and Krishnan, P. (2012). *A Framework for Analysing Driver Interactions with Semi-Autonomous Vehicles*. Paper presented at the First International Workshop on Formal Techniques for Safety-Critical Systems (FTSCS 2012), November 12, Kyoto, Japan.

Shepherd, B.E., Gilbert, P.B., and Mehrotra, D. (2007). Eliciting a counterfactual sensitivity parameter. *The American Statistician*, *61*(1), 56-63.

Short, J., Boyle, L., Shackleford, S., Inderbitzen, B., and Bergoffen, G. (2007). *The Role of Safety Culture in Preventing Commercial Motor Vehicle Crashes* (Commercial Truck and Bus Safety Synthesis 14). Washington, DC: Transportation Research Board.

Sieber, W.K., Robinson, C.F., Birdsey, J., Chen, G.X., Hitchcock, E.M., Lincoln, J.E., Nakata, A., and Sweeney, M.H. (2014). Obesity and other risk factors: The national survey of U.S. long-haul truck driver health and injury. *American Journal of Industrial Medicine*, *57*(6), 615-626.

Small, D., and Rosenbaum, P.R. (2008). War and wages: The strength of instrumental variables and their sensitivity to unobserved biases. *Journal of the American Statistical Association*, *103*(483), 924-933.

Smiley, A., Smahel. T., Boivin, D., Boudreau, P., Remmers, J., Turner, M., Rosekind, M.R., and Gregory, K.B. (2009). *Effects of a Fatigue Management Program on Fatigue in the Commercial Motor Carrier Industry*. Toronto, ON: Human Factors North, Inc.

Smith, E.B., and Ivan, J.N. (2005). *Evaluation of Safety Benefits and Potential Crash Migration Due to Shoulder Rumble Strip Installation on Freeways in Connecticut*. TRB 84th Annual Meeting Compendium of Papers CD-ROM, Washington, DC.

Smolensky, M.H., DiMilia, L., Ohayon, M.M., and Philip, P. (2011). Sleep disorders, medical conditions, and road accident risk. *Accident Analysis and Prevention*, *43*(2), 533-548.

Snyder, Q. (2013). *A Primer on Obstructive Sleep Apnea*. Washington, DC: Air Line Pilots Association, International.

Soccolich, S.A., Blanco, M., Hanowski, R.J., Olson, R.L., Morgan, J.F., Guo, F., and Wu, S.-C. (2013). An analysis of driving and working hour on commercial motor vehicle driver safety using naturalistic data collection. *Accident Analysis and Prevention*, *58*, 249-258.

Spaeth, A.M., Goel, N., and Dinges, D.F. (2014). The cumulative neurobehavioral and physiological effects of chronic caffeine intake: Individual differences and implications for the use of caffeinated energy products. *Nutrition Reviews*, *72*(Suppl. 1), 34-47.

SRF Consulting Group, Inc. (2007). *Interstate Highway Safety Study: Analysis of Vehicle Crashes Related to Safety Rest Area Spacing*. Duluth, MN: Minnesota Department of Transportation.

Starnes, M. (2006). *Large-Truck Crash Causation Study: An Initial Overview* (DOT HS 810 646). Washington, DC: National Highway Traffic Safety Administration, National Center for Statistics and Analysis.

Stein, H., and Jones, I. (1988). Crash involvement of large trucks by configuration: A case-control study. *American Journal of Public Health*, *78*(5), 491-498.

Stepanski, E.J., and Wyatt, J.K. (2003). Use of sleep hygiene in the treatment of insomnia. *Sleep Medicine Reviews*, *7*(3), 215-225.

Stevenson, M., Grunstein, R.R., Wong, K., Sharwood, L., and Elkington, J. (2014). The role of sleepiness, sleep disorders, and the work environment on heavy-vehicle crashes in two Australian states. *American Journal of Epidemiology, 179*(5), 594-601. doi 10.1093/aje/kwt305. Available: http://aje.oxfordjournals.org/content/179/5/594.full.pdf+html [April 2016].

Strine, T.W., and Chapman, D.P. (2005). Associations of frequent sleep insufficiency with health-related quality of life and health behaviors. *Sleep Medicine*, *6*(1), 23-27.

Stuart, E. (2010). Matching methods for causal inference: A review and a look forward. *Statistical Science*, *25*(1), 1-21.

Stuart, E.A., and Rubin, D.B. (2008). Matching with multiple control groups and adjusting for group differences. *Journal of Educational and Behavioral Statistics, 33*(3), 279-306.

Surani, S.R. (2014). Diabetes, sleep apnea, obesity and cardiovascular disease: Why not address them together? *World Journal of Diabetes*, *5*(3), 381-384.

Tarko, A.P. (2012). Use of crash surrogates and exceedance statistics to estimate road safety. *Accident Analysis and Prevention*, *45*, 230-240.

Taylor, W.C., Sung, N., Kolody, K.A., and Jawad, A. (1999). *A Study of Highway Rest Area Characteristics and Fatigue Related Truck Crashes*. East Lansing, MI: Michigan State University, Civil and Environmental Engineering.

Tefft, B.C. (2010). *Asleep at the Wheel: The Prevalence and Impact of Drowsy Driving*. Washington, DC: AAA Foundation for Traffic Safety.

Tefft, B.C. (2014). *Prevalence of Motor Vehicle Crashes Involving Drowsy Drivers, United States, 2009-2013*. Washington, DC: AAA Foundation for Traffic Safety.

Teran-Santos, J., Jimenez-Gomez, A., and Cordero-Guevara, J. (1999). The association between sleep apnea and the risk of traffic accidents. Cooperative Group Burgos-Santander. *New England Journal of Medicine*, *340*(11), 847-851.

Thiese, M.S., Moffitt, G., Hanowski, R.J., Kales, S.N., Porter, R.J., and Hegmann, K.T. (2015a). Commercial driver medical examinations: Prevalence of obesity, comorbidities, and certification outcomes. *Journal of Occupational and Environmental Medicine*, *57*(6), 659-665.

Thiese, M.S., Moffitt, G., Hanowski, R.J., Kales, S.N., Porter, R.J., and Hegmann, K.T. (2015b). Repeated cross-sectional assessment of commercial truck driver health. *Journal of Occupational and Environmental Medicine*, *6*(2), 104-112.

Thiffault, P. (2011). *Addressing Human Factors in the Motor Carrier Industry in Canada*. Ottawa, ON: Canadian Council of Motor Transport Administrators. Available: http://www.bv.transports.gouv.qc.ca/mono/1081534.pdf [March 2016].

Thiffault, P., and Bergeron, J. (2003). Monotony of road environment and drivers fatigue: A simulator study. *Accident Analysis and Prevention*, *35*(3), 381-391.

Thompson, J., Newman, S., and Stevenson, M. (2015). A model for exploring the relationship between payment structures, fatigue, crash risk, and regulatory response in a heavy-vehicle transportation system. *Transportation Research Part A: Policy and Practice, 82*, 204-215.

Tilley, A.J., Wilkinson, R.T., Warren, P.S.G., Watson, B., and Drud, M. (1982). The sleep and performance of shift workers. *Human Factors*, *24*(6), 629-641.

Tregear, S., Reston, J., Schoelles, K., and Phillips, B. (2009a). Continuous positive airway pressure reduces risk of motor vehicle crash among drivers with obstructive sleep apnea: Systematic review and meta-analysis. *Sleep*, *33*(10), 1373-1380.

Tregear, S., Reston, J., Schoelles, K., and Phillips, B. (2009b). Obstructive sleep apnea and risk of motor vehicle crash: Systematic review and meta-analysis. *Journal of Clinical Sleep Medicine*, *5*(6), 573-581.

Turino, G.M., Bernstein, L., Dauber, J.H., George, R.D., Golding, R.M., Hudson, L.D., Kin, T.E., Manhay, R.A., Maynock, T.L., Panerson, R., Richerson, H.B., Rodarte, J.R., Scharer, L., Stover, D.E., and Zwillich, C.W. (1991). *Conference on Pulmonary/Respiratory Disorders and Commercial Drivers* (FHWA-MC-91-004). Available: https://www.fmcsa.dot.gov/regulations/medical/conference-pulmonaryrespiratory-disorders-and-commercial-drivers [March 2016].

U.S. Bureau of Labor Statistics. (2014). *Heavy and Tractor-Trailer Truck Drivers*. Available: http://www.bls.gov/ooh/transportation-and-material-moving/heavy-and-tractor-trailer-truck-drivers.htm [March 2016].

U.S. Department of Transportation. (2013). *Large Truck and Bus Crash Facts 2013* (FMCSA-RRA-15-004). Washington, DC: U.S. Department of Transportation, Federal Motor Carrier Safety Administration.

U.S. Department of Transportation. (2014). *Passenger Travel Facts and Figures 2014*. Available: http://www.rita.dot.gov/bts/publications/passenger_travel [March 2016].

U.S. Federal Highway Administration. (1996). *Commercial Driver Rest & Parking Requirements: Making Space for Safety* (FHWA-MC-96-0010). Washington, DC: U.S. Department of Transportation, U.S. Federal Highway Administration.

U.S. Federal Highway Administration. (2013). *Highway Statistics 2013*. Available: http://www.fhwa.dot.gov/policyinformation/statistics/2013 [March 2016].

Van Dongen, H.P.A., and Belenky, G. (2010). *Investigation into Motor Carrier Practices to Achieve Optimal Commercial Motor Vehicle Driver Performance: Phase I* (FMCSA-RRR-10-005). Washington, DC: U.S. Department of Transportation, Federal Motor Carrier Safety Administration.

Van Dongen, H.P.A., and Mollicone, D.J. (2014). *Field Study on the Efficacy of the New Restart Provision for Hours of Service* (RRR-13-058). Washington, DC: U.S. Department of Transportation, Federal Motor Carrier Safety Administration.

Van Dongen, H.P.A., Price, N.J., Mullington, J.M., Szuba, M.P., Kapoor, S.C., and Dinges, D.F. (2001). Caffeine eliminates psychomotor vigilance deficits from sleep inertia. *Sleep*, *24*(7), 813-819.

Van Dongen, H.P.A., Maislin, G., Mullington, J.M., and Dinges, D.F. (2003a). The cumulative cost of additional wakefulness: Dose-response effects on neurobehavioral functions and sleep physiology from chronic sleep restriction and total sleep deprivation. *Sleep*, *26*(2), 117-126.

Van Dongen, H.P.A., Rogers, N.L., and Dinges, D.F. (2003b). Understanding sleep debt: Theoretical and empirical issues. *Sleep and Biological Rhythms*, *1*, 5-13.

Van Dongen, H.P.A., Mott, C.G., Huang, J-K, Mollicone, D.J., McKenzie, F.D., and Dinges, D.F. (2007). Optimization of biomathematical model predictions for cognitive performance impairment in individuals: Accounting for unknown traits and uncertain states in homeostatic and circadian processes. *Sleep*, *30*(9), 1125-1139.

Victor, T., Bärgman, J., Boda, C., Dozza, M., Engstroem, J., Flannagan, C., Lee, J.D., and Markkula, G. (2014). *Analysis of Naturalistic Driving Study Data: Safer Glances, Driver Inattention, and Crash Risk* (S2-S08A-RW-1). Washington, DC: Transportation Research Board, National Academy of Sciences.

Victor, T., Dozza, M., Bärgman, J., Boda, C.-N., Engström, J., Flannagan, C., and Markkula, G. (2015). *Analysis of Naturalistic Driving Study Data: SAFER Glances, Driver Inattention, and Crash Risk* (SHRP Safety Project SO8). Washington, DC: Transportation Research Board, National Academy of Sciences.

Watson, N.F., Badr, M.S., Belenky, G., Bliwise, D.L., Buxton, O.M., Buysse, D., Dinges, D.F., Gangwisch, J., Grandner, M.A., Kushida, C., Malhotra, R.K., Martin, J.L., Patel, S.R., Quan, S.F., and Tasali, E. (2015a). Joint consensus statement of the American Academy of Sleep Medicine and Sleep Research Society on the recommended amount of sleep for a healthy adult: Methodology and discussion. *Sleep*, *38*(8), 1161-1183.

Watson, N.F., Badr, M.S., Belenky, G., Bliwise, D.L., Buxton, O.M., Buysse, D., Dinges, D.F., Gangwisch, J., Grandner, M.A., Kushida, C., Malhotra, R.K., Martin, J.L., Patel, S.R., Quan, S.F., and Tasali, E. (2015b). Recommended amount of sleep for a healthy adult: A joint consensus statement of the American Academy of Sleep Medicine and Sleep Research Society. *Sleep*, *38*(6), 843-844.

Weaver, T.E., Laizner, A.M., Evans, L.K., Maislin, G., Chugh, D.K., Lyon, K., Smith, P.L., Schwarta, A.R., Redline, S., Pack, A.I., and Dinges, D.F. (1997). An instrument to measure functional status outcomes for disorders of excessive sleepiness. *Sleep*, *20*(10), 835-843.

West, S.G., Duan, N., Pequegnat, W., Gaist, P., Des Jarlais, D.C., Holtgrave, D., Szapocznik, J., Fishbein, M., Rapkin, B., Clatts, M., and Dolan Mullen, P. (2008). Alternatives to the randomized controlled trial. *American Journal of Public Health*, *98*(8), 1359-1366.

Wierwille, W.W., and Ellsworth, L.A. (1994). Evaluation of driver drowsiness by trained raters. *Accident Analysis and Prevention*, *26*(5), 571-581.

Wierwille, W.W., Wreggit, S.S., Kim, C.L., Ellsworth, L.A., and Fairbanks, R.J. (1994). *Research on Vehicle-based Driver Status/Performance Monitoring: Development, Validation, and Refinement of Algorithms for Detection of Driver Drowsiness* (DOT HS 808 247). Washington, DC: National Highway Traffic Safety Administration.

Williams, J.R., Amana, A., and Tregear, S.J. (2011). *Obstructive Sleep Apnea and Commercial Motor Vehicle Driver Safety: Updated Review*. Available: http://ntl.bts.gov/lib/44000/44600/44638/OSA_Update_11302011-p.pdf [March 2016].

Williams, V.H., and McLaughlin, S.B. (2013). *A Survey of Motorcyclists: Data for Research Design and Instrumentation* (Report 13-UT-019). Blacksburg, VA: Virginia Tech Transportation Institute, National Surface Transportation Safety Center for Excellence.

Williamson, A., and Friswell, R. (2013). The effect of external non-driving factors, payment type and waiting and queuing on fatigue in long distance trucking. *Accident Analysis and Prevention, 58*(Sept.), 26-34. doi: 10.1016/j.aap.2013.04.017.

Williamson, A., Lombardi, D.A., Folkard, S., Stutts, J., Courtney, T.K., and Connor, J.L. (2011). The link between fatigue and safety. *Accident Analysis and Prevention*, *43*(2), 498-515.

Wittes, J., Lakatos, E., and Probstfield, J. (1989). Surrogate endpoints in clinical trials: Cardiovascular diseases. *Statistics in Medicine*, *8*(4), 415-425.

Wood, J.S., Gooch, J.P., and Donnell, E.T. (2015). Estimating the safety effects of lane widths on urban streets in Nebraska using the propensity scores-potential outcomes framework. *Accident Analysis and Prevention*, *82*, 180-191.

Wu, H., and Yan-Go, F. (1996). Self-reported automobile accidents involving patients with obstructive sleep apnea. *Neurology*, *46*(5), 1254-1257.

Wylie, C., Shultz, T., Miller, J., Mitler, M., and Mackie, R. (1996). *Commercial Motor Vehicle Driver Fatigue and Alertness Study: Project Report* (FHWA-MC-97-002). Washington, DC: U.S. Department of Transportation, Federal Highway Administration.

Xie, L., Kang, H., Xu, Q., Chen, M.J., Liao, Y., Thiyagarajan, M., O'Donnell, J., Christensen, D.J, Nicholson, C., Iliff, J.J., Takano, T., Deane, R., and Nedergaard, M. (2013). Sleep drives metabolite clearance from the adult brain. *Science*, *342*(6156), 373-377.

Yamamoto, H., Akashiba, T., Kosaka, N., Ito, D., and Horie, T. (2000). Long-term effects nasal continuous positive airway pressure on daytime sleepiness, mood and traffic accidents in patients with obstructive sleep apnoea. *Respiratory Medicine*, *94*(1), 87-90.

Yang, G., Lai, C.S., Cichon, J., Ma, L., Li, W., and Gan, W.B. (2014). Sleep promotes branch-specific formation of dendritic spines after learning. *Science*, *344*(6188), 1173-1178.

Young, T., Evans, L., Finn, L., and Palta, M. (1997). Estimation of the clinically diagnosed proportion of sleep apnea syndrome in middle-aged men and women. *Sleep*, *20*(9), 705-706.

Zhang, H., Yan, X., Wu, C., and Qiu, T. (2014). Effect of circadian rhythms and driving duration on fatigue level and driving performance of professional drivers. *Transportation Research Record: Journal of the Transportation Research Board*, *2402*, 19-27.

Zubizarreta, J.R., Cerda, M., and Rosenbaum, P.R. (2013). Designing an observational study to be less sensitive to unmeasured biases: Effect of the 2010 Chilean earthquake on posttraumatic stress. *Epidemiology*, *24*, 79-87.

Appendix

Biographical Sketches of Panel Members and Staff

MATTHEW RIZZO *(Cochair)* is the Francis and Edgar Reynolds chair and professor of the Department of Neurological Sciences at the University of Nebraska Medical Center and codirector of the Nebraska Neuroscience Alliance. Previously, he was a senior faculty member in the Department of Neurology and the Division of Behavioral Neurology and Cognitive Neuroscience at the University of Iowa where he was also director of the Visual Functional Laboratory, Simulator for Research in Ergonomics and Safety, Automobile for Research in Ergonomics and Safety, and Nissan-Iowa Instrumented Vehicles for Ergonomics and Neuroscience. His long-term research and work has focused on the association between human factors and the performance of operators of motor vehicles. He is a fellow of the American Neurological Association and the American Academy of Neurology. He has been honored with the Inaugural Medical Review Board Service Award of the U.S. Department of Transportation. He has an A.B. from Columbia University and an M.D. from the Johns Hopkins University School of Medicine.

HAL S. STERN *(Cochair)* is the Ted and Janice Smith Family Foundation dean and professor of statistics in the Donald Bren School of Information and Computer Sciences at the University of California, Irvine. Previously, he held faculty positions and served as director of undergraduate studies in the department of statistics at Harvard University. He also previously held the Laurence H. Baker chair in biological sciences and served as director of graduate studies at Iowa State University. His primary areas of research are Bayesian methods, model diagnostics, and statistical applications to the biological and social sciences. He is a fellow of the Institute of Mathematical Statistics and the American Statistical Association. He has a B.S. in mathematics from Massachusetts Institute of Technology and an M.S. and a Ph.D. in statistics from Stanford University.

DANIEL BLOWER is an associate research scientist with the Center for the Management of Information for Safe and Sustainable Transportation at the University of Michigan Transportation Research Institute. His primary area of research is traffic crash causation, focusing primarily on medium and heavy trucks, and he has also directed projects on traffic safety issues related to light vehicles. His past projects have included investigating the crash experience of younger truck drivers, developing an event tree for heavy truck accidents, and developing statistical models relating vehicle configuration and operating environment to the probability of accident involvement. He is chair of the Michigan Truck Safety Commission, and he previously served on the Technical Advisory Group for American Transportation Research Institute on Truck Drivers Hours of Service study; the Technical Advisory Committee on National Motor Vehicle Crash Causation Study for the National Highway Traffic Safety Administration; and the Large Truck Crash Causation Study committee for the Federal Motor Carrier Safety Administration. He has a B.A. and a Ph.D. in history from the University of Michigan.

MICHAEL L. COHEN (*Costudy Director*) is a senior program officer for the Committee on National Statistics. He has led or served as contributing staff on a wide range of studies on the U.S. census and the modeling and reliability of defense systems. He also serves as a consultant on statistical analysis for other divisions in the National Academies of Sciences, Engineering, and Medicine. Previously, he was a mathematical statistician at the Energy Information Administration and held positions at the School of Public Affairs at the University of Maryland and at Princeton University. His general area of interest is the use of statistics in public policy, with particular focus in census undercount, model validation, and robust estimation. He is a fellow of the American Statistical Association and an elected member of the International Statistical Institute. He has a B.S. in mathematics from the University of Michigan and an M.S. and a Ph.D. in statistics from Stanford University.

CHARLES A. CZEISLER is the Baldino professor of sleep medicine at Harvard Medical School. His research interests include basic and applied research on the physiology of human circadian rhythm and its relationship to the sleep-wake cycle, including the application of sleep science and sleep medicine to policy on occupational medicine and health policy, particularly as it relates to extended duration work shifts and long work weeks. He is a member of the National Academy of Medicine and the American Clinical and Climatological Association, a fellow of the Clinical Sleep Society, the American Sleep Disorders Association, the Academy of Behavioral Medicine Research, and the American Society for Clinical Investigation, and an honorary fellow of the Royal College of Physicians. He is a recipient of the Lifetime Achievement Award of the National Sleep Foundation, the Lord Adrian Gold Medal from the Royal Society of Medicine, and the distinguished scientist award from the Sleep Research Society. He has an M.D. from the Stanford University School of Medicine and a Ph.D. in neuro- and biobehavioral sciences from Stanford University.

DAVID F. DINGES is a professor and vice chair in the Department of Psychiatry at the University of Pennsylvania Perelman School of Medicine, where he also serves as chief of the Division of Sleep and Chronobiology and director of the Unit for Experimental Psychiatry. He also heads the Neurobehavioral and Psychosocial Factors Team for the National Space Biomedical Research Institute. His research has focused on laboratory and field studies of the physiological, neurobehavioral, and fatiguing effects of inadequate sleep in humans; the causes and consequences for accidents and catastrophic events associated with human error; and the behaviors, countermeasures and technologies that can prevent or mitigate the effects of fatigue on human safety. He is a fellow of the Academy of Behavioral Medicine Research, the American Physiological Society, the American Psychological Association, and the Association for Psychological Science, an overseas fellow of the International Association of Traffic and Safety Societies, and an elected member of the International Academy of Astronautics. He is a recipient of the Decade of Behavior Research Award from the American Psychological Association and the Distinguished Public Service Medal from the U.S. National Aeronautics and Space Administration. He has an A.B. from Benedictine College and an M.S. and a Ph.D. in physiological psychology from Saint Louis University.

JOEL B. GREENHOUSE is a professor in the department of statistics at Carnegie Mellon University and an adjunct professor of psychiatry and adjunct professor of epidemiology at the University of Pittsburgh. Previously, he served as associate dean for academic affairs and as assistant professor in the Department of Statistics at Carnegie Mellon University. His research focuses on biostatistical applications and meta-analysis. He is an elected fellow of the American Statistical Association, the American Association for the Advancement of Science, and the International Statistical Institute. He is a recipient of the Statistician of the Year Award from the Pittsburgh Chapter of the American Statistical Association. He has a B.S. in mathematics from the University of Maryland and an A.M. in statistics and an M.P.H. and a Ph.D. in biostatistics from the University of Michigan.

FENG GUO is an associate professor in the Department of Statistics at Virginia Polytechnic Institute and State University, where he is affiliated with the Virginia Tech Transportation Institute. His primary areas of research are transportation statistics, Bayesian hierarchical models, traffic safety models, spatial statistics, and statistical epidemiology. His specific area of interest is driver behavior and analysis of data from naturalistic driving studies. He is a member of the Committee on Statistical Methods and Committee on Safety, Data, Analysis, and Evaluation of the Transportation Research Board of the National Academies. He has a B.S. in highway and traffic engineering and an M.S. in transportation economics and management from Tongju University in Shanghai, China, and Ph.D. degrees in transportation engineering and statistics from the University of Connecticut.

RICHARD J. HANOWSKI is director of the Center for Truck and Bus Safety and senior research scientist for the Virginia Tech Transportation Institute. His research focuses on transportation human factors involving both heavy and light vehicles, laboratory and field testing, simulation, advanced system development and testing, naturalistic driving, and human performance evaluation. His area of expertise is fatigue and distracted driving, particularly pertaining to commercial vehicle drivers. He is a recipient of the Paul S. Richards Endowed Distinguished Visiting Lectureship in Occupational Health, the 2011 SAE International L. Ray Buckendale Lecture Award, and was the recipient of HFES's Best *Ergonomics in Design* Article Award. He has a B.A. in psychology from the University of Saskatchewan in Saskatoon, Canada, an M.S. in psychology from the University of Idaho, and a Ph.D. in industrial and systems engineering (human factors engineering) from Virginia Polytechnic Institute and State University.

NATALIE P. HARTENBAUM is president and chief medical officer of OccuMedix, Inc., an occupational medicine consulting firm that provides occupational health and safety consulting services, including expert medical review and litigation support on issues of the Americans with Disabilities Act and fitness for duty. She also serves on both the clinical and teaching faculty of the University of Pennsylvania and as the medical director for several companies in the Philadelphia area. Previously, she served as medical director of the Consolidated Railroad Corporation, associate medical director for CentraMed Occupational Health Specialists, and medical director at the Industrial Health Care Center. She is past president of the American College of Occupational and Environmental Medicine, a member of the Board of Trustees of the American Board of Preventive Medicine, and president of the Philadelphia Occupational and Environmental Medicine Society. She is a recipient of the American College of Occupational and Environmental Medicine's President's Award. She has a B.A. in biology from Temple University, an M.D. from the Temple University School of Medicine, and an M.P.H. in occupational medicine from the Medical College of Wisconsin.

GERALD P. KRUEGER is president of Krueger Ergonomics Consultants in Alexandria, Virginia. His primary research interests are human factors, human engineering, ergonomics, operator fatigue, and employee health, wellness, and fitness. His research expertise includes the human performance implications of equipment operator fatigue, sleep deprivation, sustained operations, and formulating worker health and wellness programs. Previously, he worked as principal scientist/ergonomist and as director of human factors, ergonomics, and medical research programs for the Wexford Group. He also held positions at UTEK Corporation, Star Mountain, Inc., and the Biomechanics Corporation of America. He spent 25 years doing occupational medicine research in the U.S. Army, ultimately achieving the position of colonel, with the U.S. Army Research Institute of Environmental Medicine. He served at the U.S. Army Aeromedical Research Laboratory; the Walter Reed Army Institute of Research; the Headquarters of the U.S. Army Medical Research and Development Command; and as medical R&D command liaison officer at the U.S. Army Human Engineering Laboratory. He is a fellow of the American Psychological Association, the Human Factors and Ergonomics Society, the Institute of Ergonomics and Human Factors, the Washington Academy of Sciences, and he is an associate fellow of the Aerospace Medical Association. He has been awarded the Order of Military Medical Merit for Career Contributions to the U.S. Army Medical Department and a presidential citation for military service as a psychologist in the Vietnam conflict. In 2015 the American Psychological Association awarded him the Flannagan Award for Lifetime Achievement in Military Psychology. He served over 20 years as book review editor for Ergonomics in Design; is an associate editor for Military Psychology; served as section editor for Aviation, Space, and Environmental Medicine; and he was a member of the editorial board of the Society for Human Performance in Extreme Environments. He has a B.A. in psychology from the University of Dayton and an M.A. in engineering psychology and Ph.D. in experimental and engineering psychology from the Johns Hopkins University.

MELISSA M. MALLIS is president and chief scientist of M3Alertness Management, chief scientific advisor for Alertness Solutions, and senior science advisor for DB&A, a management consulting firm. She is also a fellow at George Mason University in the Center for Infrastructure Protection and Homeland Security. Previously, she served as chief scientist for operational and fatigue research at the Institutes for Behavioral Resources and as director of scientific affairs at Alertness Solutions. Her major areas of interest are the development of innovative, practical, and effective strategies to enhance safety, performance, and alertness in various 24/7 operational environments. She has received several NASA superior performance awards and incentive awards from the U.S. National Aeronautics and Space Administration. She has also received the Arnold D. Tuttle award from the Aerospace Medical Association and is a three-time awardee of the William E. Collins award from the Aerospace Human Factors Association. She has a B.S. in physics from Villanova University and a Ph.D. in biomedical science from Drexel University.

RICHARD PAIN (*Consultant*) recently retired from the Transportation Research Board, where he was the transportation safety coordinator in the board's Technical Activities Division. He served as staff to a wide range of committees, including studies of truck and bus safety, statistics in transportation, visualization in transportation, and future truck and bus safety research opportunities, as well as an international conference on research on the health and wellness of commercial truck and bus drivers. Prior to his work for the Transportation Research Board, his work focused on human factors and safety research and evaluation in the transportation, nuclear, civil, and military areas, conducting numerous laboratory, simulation, and fully operational experiments; training, development, conduct, and evaluation studies; and human engineering reviews. He has a B.A. in psychology from Hofstra University and an M.A. in clinical psychology and a Ph.D. in applied experimental psychology from Michigan State University.

JOHN R. PEARSON is program director of the Council of Deputy Ministers Responsible for Transportation and Highway Safety in Ottawa, Ontario, where he is responsible for developing, managing, and conducting research on policy development programs. He previously served as a consultant to the Council of Deputy Ministers; executive director of the Canadian Trucking Research Institute; director of technical programs for the Transportation Association of Canada; director of research for the Western Highway Institute; and project manager for the Vehicle Weights and Dimensions Study at the Canroad Transportation Research Corporation. His research interests include safety using naturalistic driving techniques, with expertise in highway safety, especially weights and dimensions policies. He has twice received chairman's award of the Roads and Transportation Association of Canada. He has a bachelor of engineering degree from Carleton University in Ottawa, Ontario.

ESHA SINHA (*Costudy Director*) is an associate program officer for the Committee on National Statistics. She has served as a staff officer for a wide range of projects, including studies on the measurement of productivity in higher education, future content and methods for R&D resources, and, currently, on compliance, safety, and accountability for federal motor carriers. Previously, she worked for the Indian Institute of Management in Ahmedabad. At the State University of New York at Binghamton, she worked extensively on student records on such topics as whether advanced placement or SAT scores are better predictors of college success and performance of transfer students. She has an M.A. degree in economics from GIPE, India, and a Ph.D. in economics from SUNY Binghamton.

DYLAN SMALL is a professor in the Department of Statistics at the Wharton School of the University of Pennsylvania and an associate scholar in the university's Biostatistics Unit of the Center for Clinical Epidemiology and Biostatistics in the School of Medicine. His areas of research include causal inference, the design and analysis of observational studies, and applications of statistics to public health, medicine, and public policy. He is the founding editor of the journal *Observational Studies*. He is a fellow of the American Statistical Association and a senior fellow at the Leonard Davis Institute of Health Economics. He has an A.B. in mathematics from Harvard University and a Ph.D. in statistics from Stanford University.

ELIZABETH A. STUART is a professor in the Department of Mental Health, the Department of Biostatistics, and the Department of Health Policy and Management in the Johns Hopkins Bloomberg School of Public Health. Previously, she was a researcher at Mathematica Policy Research. Her primary areas of research include statistical methodology for mental health research, particularly relating to causal inference and missing data. She also conducts research in the areas of education, prevention, and intervention using techniques such as multilevel modeling, matching, and propensity scores. She is a fellow of the American Statistical Association. She has an A.B. in mathematics and chemistry from Smith College and an A.M. and a Ph.D. in statistics from Harvard University.

DAVID H. WEGMAN is professor emeritus and founding chair of the Department of Work Environment of the School of Health and Environment at the University of Massachusetts, Lowell. He is also an adjunct professor at the Harvard School of Public Health and vice president of the Alpha Foundation for the Improvement of Mine Safety and Health. His epidemiologic research includes the study of acute and chronic occupational respiratory disease, cancer risk, and musculoskeletal disorders, and he also studies the subjective outcomes as early indicators of health effects, surveillance of occupational conditions and risks, and occupational health policy. He chaired the Mine Safety and Health Administration Advisory Committee on the Elimination of Pneumoconiosis Among Coal Mine Workers for the U.S. Department of Labor, the International Evaluation Group for an analysis of Occupational Health Research in Sweden, and the United Auto Workers/General Motors Occupational Health Advisory Board. He is a fellow of the American College of Epidemiology, and he is a recipient of the Alice Hamilton Lifetime Achievement Award in the Occupational Health and Safety Section from the American Public Health Association. He has also served as chair of He has a B.A. in history from Swarthmore College and an M.Sc. in occupational health and an M.D. from Harvard University.

COMMITTEE ON NATIONAL STATISTICS

The Committee on National Statistics was established in 1972 at the National Academies of Sciences, Engineering, and Medicine to improve the statistical methods and information on which public policy decisions are based. The committee carries out studies, workshops, and other activities to foster better measures and fuller understanding of the economy, the environment, public health, crime, education, immigration, poverty, welfare, and other public policy issues. It also evaluates ongoing statistical programs and tracks the statistical policy and coordinating activities of the federal government, serving a unique role at the intersection of statistics and public policy. The committee's work is supported by a consortium of federal agencies through a National Science Foundation grant.