

# STATISTICAL ISSUES ON EVALUATING FATIGUE AND DRIVING RISK

**10<sup>th</sup> International Conference on Managing Fatigue**

Feng Guo

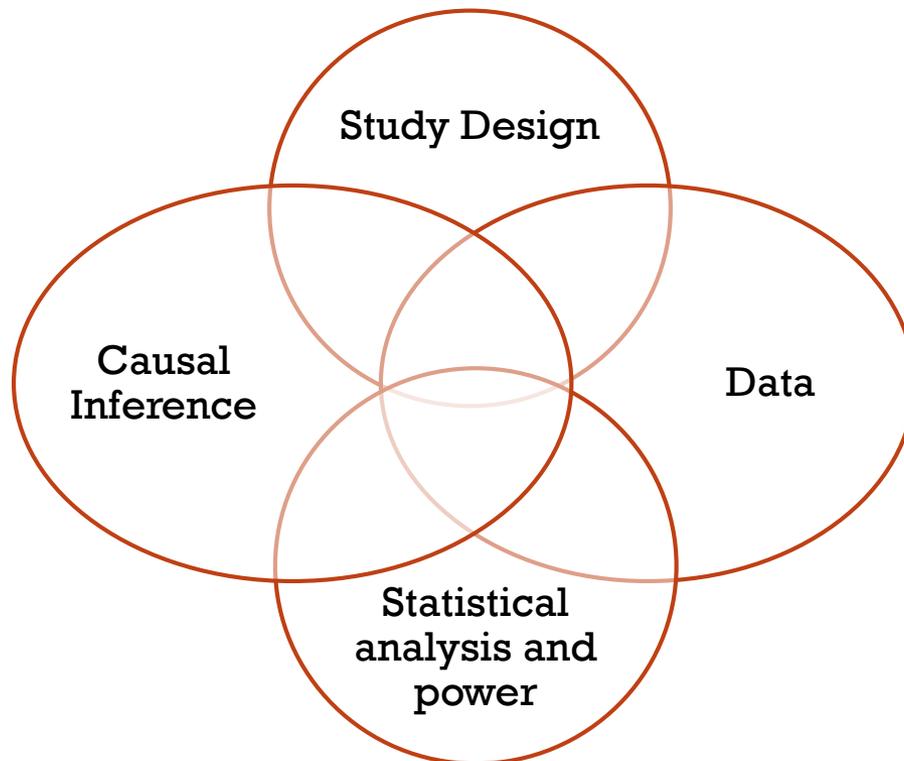
San Diego, CA

March 21, 2017



# EVALUATE THE RELATIONSHIP BETWEEN FATIGUE AND DRIVING SAFETY

- Objective: to establish causal relationship



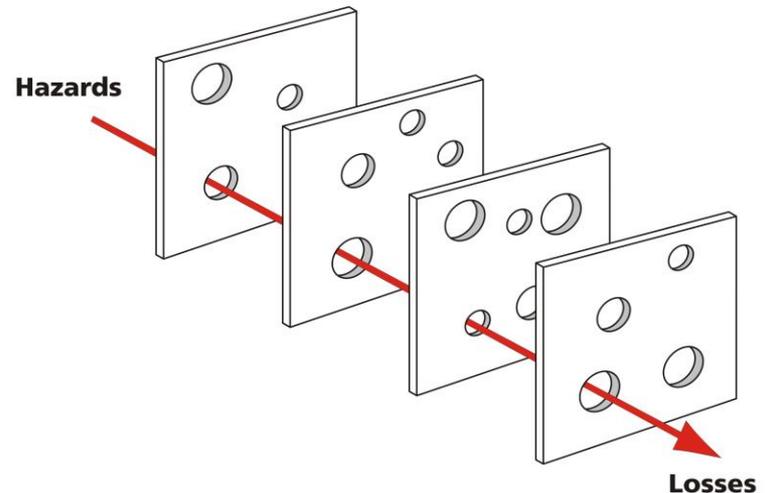
# STUDY DESIGN

- ❖ Define overall structure of the study
- ❖ Guide how the data should be collected and analyzed
- ❖ How strong the evidence toward causal conclusion
  
- Randomized controlled trials
  - Exposure assigned by researchers
  - Better control for confounding
  - Lab or simulator studies
  - May not be feasible or ethical for some studies
  
- Observational studies
  - Exposure status self-selected by participants
  - Generally feasible
  - More susceptible for to confounding
  - Cohort; case-control; case-crossover; cross-sectional designs



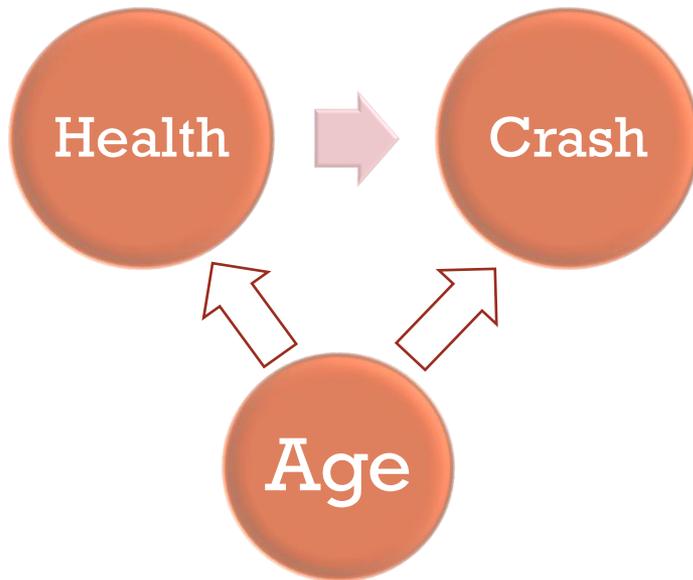
# CHALLENGES IN CAUSAL INFERENCE

- Complex causal structure
  - Vehicle
  - Driver
  - Fleet
  - Traffic/Environment
- Models
  - Swiss cheese model
  - Haddon matrix
  - Minimum sufficient risk set
- Causal effect
  - Potential outcomes
  - Association  $\neq$  causal effect

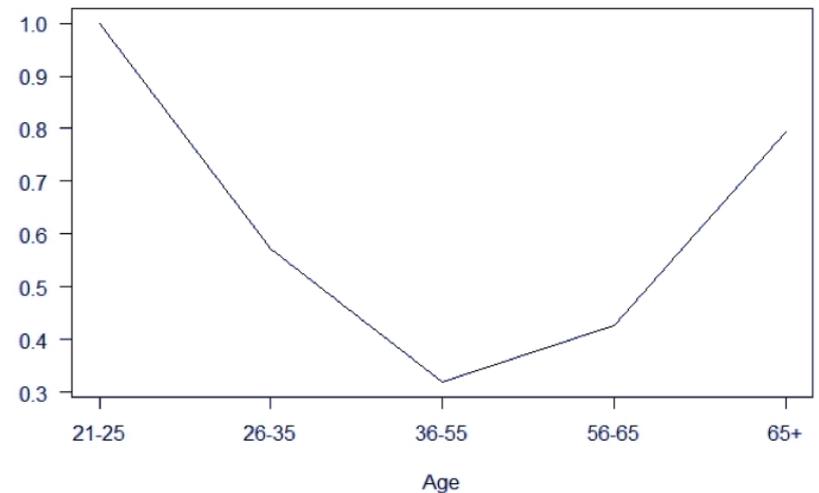


# ANALYSIS METHODS FOR OBSERVATIONAL STUDIES

- Confounding: other risk factors leading to bias in fatigue-crash relationship



Relative Crash Risk by Age



# ANALYSIS OBSERVATIONAL STUDIES

- Control through design
  - Matched cohort or case-control
  - Case-crossover
- Analysis options
  - Regression
  - Propensity score
  - Instrumental variables
  - Regression discontinuity
  - Marginal structure models
  - ...
- Limited use in traffic safe analysis.
  - Data: “Err on the side of including variables whose status regarding causal impact on crash risk is unclear.”



# ESTABLISH CAUSAL RELATIONSHIP: HILL'S CRITERIA

- Strength
- Temporal relationship
- Consistency\*
- Theoretical plausibility\*
- Coherence\*
- Specificity
- Dose-response relationship
- Experimental evidence\*
- Analogy

\* Elimination of alternative explanations

\* Multiple studies/sources

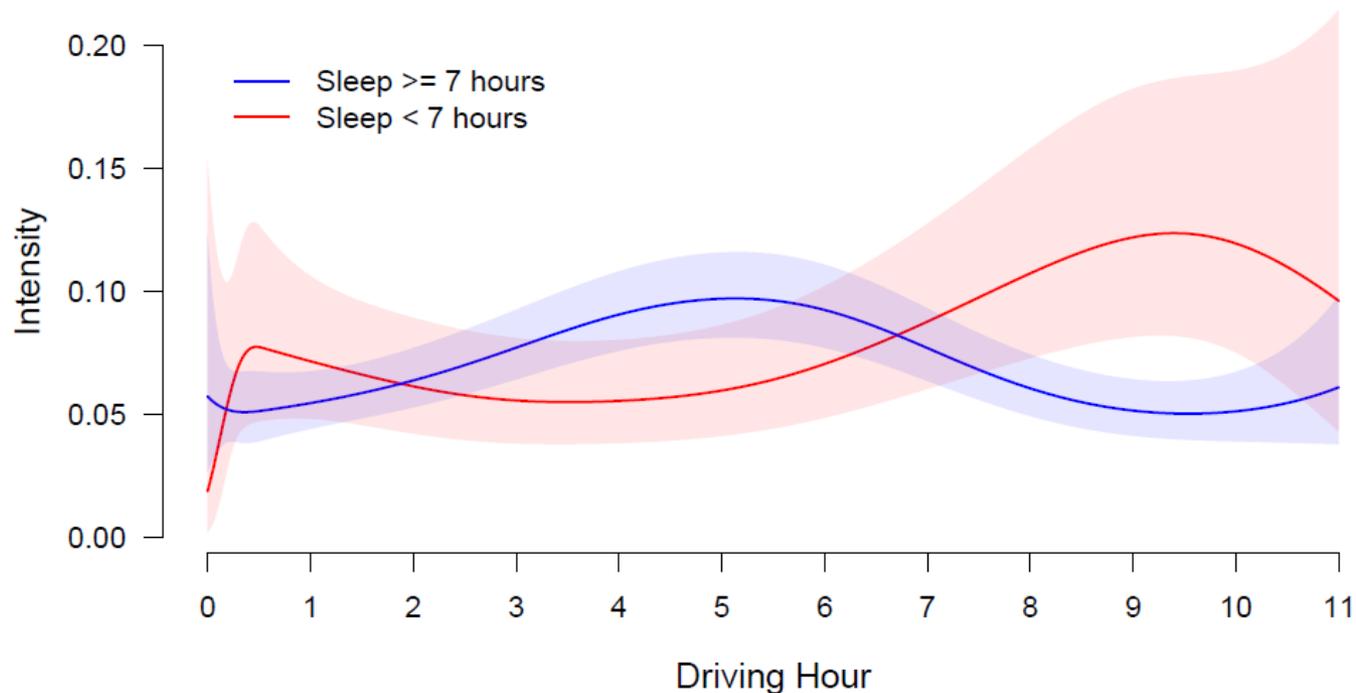


# STATISTICAL METHODS

- Depending on corresponding study design
- Generalized linear models: Poisson/Negative Binomial, logistics regression
- Correlation among observation is critical, especially for high fidelity data with multiple observations from same driver/fleet.
  
- Statistical power
  - Low exposure and sample size for long trips (e.g., exposure on driving over 11<sup>th</sup> hour)
  - Type I and Type II error: alternative hypothesis difficult to define
  - Statistical significance vs. clinically meaningful difference



# RECURRENT EVENT MODEL FOR DRIVING PERFORMANCE USING NDS DATA



**Figure:** The estimated intensity functions (in solid line),  $\hat{\lambda}(t|sleep = 0)$  and  $\hat{\lambda}(t|sleep = 1)$ , with 95% pointwise confidence intervals (in shaded area).



# BUILD TRANSPORTATION STATISTICS COMMUNITY

- **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.
- The first Transportation Statistics Summer Workshop to be held at the Statistical and Applied Mathematical Sciences Institute August, 2017
- The Transportation Statistics Interest Group of American Statistical Association is filing a Petition to become a section.

