Time-Series Modeling in Driving Studies

Tenecia Harris
Dr. Gideon Zamba

Department of Biostatistics
University of Iowa

Outline

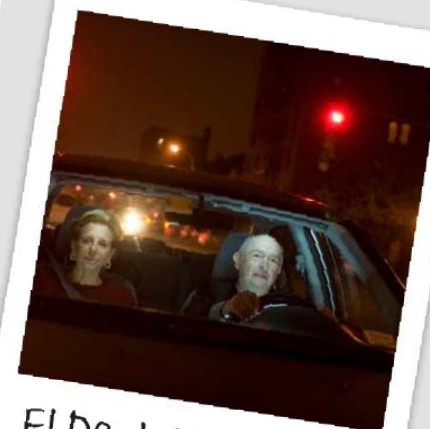
- Introduction
- Proposed Model
- Application of Model
- Results/Analysis
- Conclusion/Future Work
- Acknowledgements

Introduction

Talking and Driving!

Radio Traffic announcements

GPS



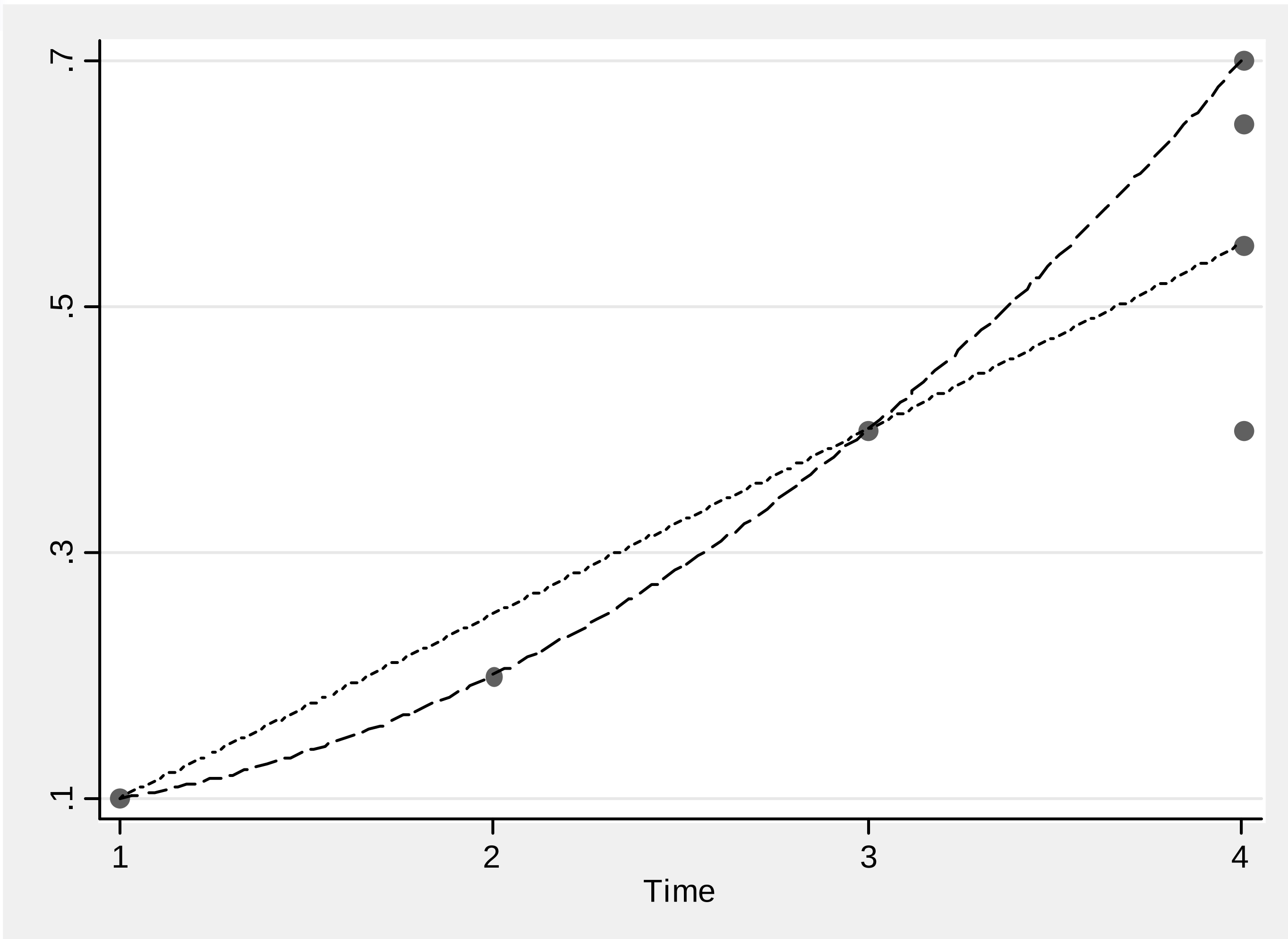
ELDERLY Drivers

- Lateral control- the nature of control on the positioning of the vehicle
    - How is lateral control measured?
        - Driving simulators coupled with evaluation models/methods

Proposed Model

- Third-order time-series
- Lane position at time *t* can be predicted by the previous 3 time points
  - Let θ_t denote the lateral position at time t. For t>3
  - $\theta_t = g(\theta_{t-1}, \theta_{t-2}, \theta_{t-3}) + |e_t|I_t$
    - g( ) : an unknown function which predicts the lateral position at time *t* based on the observed previous three time points. e_t and I_t are model arguments to which we will return

Projections based on previous points



Continuation of the model… Reparameterization

$\theta'_{t1} = \theta_{t-1}$ (Flat projection)

$\theta'_{t2} = \theta_{t-1} + (\theta_{t-1} - \theta_{t-3})/2$ (Linear interpolation)

$\theta'_{t3} = 3\theta_{t-1} - 3\theta_{t-2} + \theta_{t-3}$ (Quadratic projection)

Each of these projections individually doesn't estimate lateral control.

A convex combination of them does.

Continuation of proposed model… Reparameterization and assumptions

$$E(\theta_t) = \beta_1 \theta'_{t1} + \beta_2 \theta'_{t2} + \beta_3 \theta'_{t3}$$

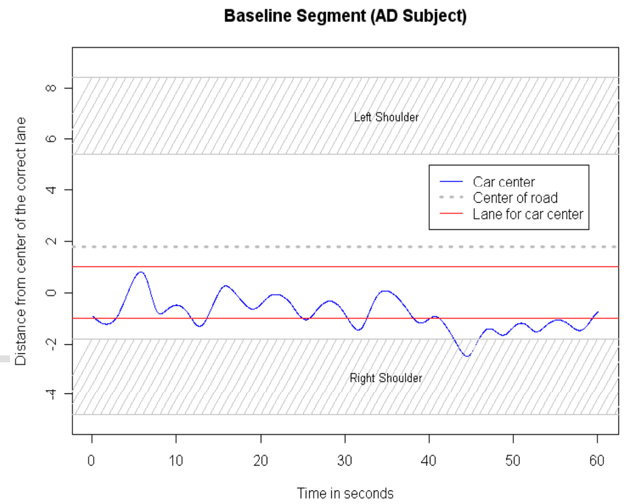
$$0 \leq (\beta_1, \beta_2, \beta_3) \leq 1 \text{ ; and } \beta_1 + \beta_2 + \beta_3 = 1$$

This parameterization, although taking into account all three projections doesn't incorporate the idea of centering or re-centering a vehicle if necessary. Reason for the following model:

$$\theta_t = \beta_1 \theta'_{t1} + \beta_2 \theta'_{t2} + \beta_3 \theta'_{t3} + |e_t| I_t$$

Where $|e_t| I_t$ is closely connected to re-centering the vehicle.

Cont.



- $e_t \sim N(0,\sigma^2)$
  - e_t: normally distributed error between observed and predicted position at time t ;σ^2 the error variance
  - I_t : a sign indicator, equaling -1 and 1 with probability p_t and $1-p_t$, respectively
  - This is clearly connected to real world when a person chooses to re-center a vehicle in case poor lateral control occurs. This will be helpful if we were to use our method in say GPS, grading driver, warning drivers under workload, keeping safe drivers on the road, assisting neurologically affected drivers etc…

Continuation of proposed model

- Re-centering at time *t* comes about when there seemed to be a poor lateral control at time *t-1.* So the probability of re-centering is based on knowledge of the preceding position.
- If at time *t* the observed is greater than the predicted, it results in positive residual; otherwise, the residual is negative. This yields modeling the sign of the residuals as a logistic function
- $log\ (p_t /([1-p_t]) = \gamma_0 + \gamma_1\ \theta_{t-1}$
  - γ_0: The intercept of the logistic model. Exp(γ_0) describes the odds of having positive residuals at time *t* when the driver's previous point is already at the center.
  - γ_1: The slope of the logistic model. Exp(γ_1) is the augmentation in the odds of positive residual when lateral control at time *t-1* increases by 1
  - θ_{t-1}: The observed lateral position at the previous time point.

SIREN --Simulator for Interdisciplinary Research in Ergonomics and Neuroscience

Application of model from a previous study

- 67 subjects with Alzheimer’s disease and 128 elderly neurologically normal subjects took a driving test and their lane position was recorded.
- We use these recording to estimate the parameters of our model and how the parameters discriminate these two neurological groups of subjects.
- The following table gives to results

Estimated parameters

Parameter being estimated, or variability measure	Mean (SD)		T-test Statistics (AD vs. Normal)
	Normal	AD	
β_1 (Flat)	.055 (.018)	0.052 (.020)	-1.07
β_2 (Linear)	0.47 (0.25)	0.31 (0.24)	-4.22*
β_3 (Quadratic)	0.48 (0.25)	0.64 (0.24)	4.37*
σ^2	0.0046 (0.0008)	0.0046 (0.0013)	-0.39
γ_0	0.63(0.79)	0.42(0.54)	2.02
γ_1 (Re-centering)	2.29 (1.35)	1.63 (1.14)	-3.44*
Entropy	0.60 (0.03)	0.59 (0.03)	-1.48
SD of Prediction Error	0.0145 (0.0056)	0.0150 (0.0068)	0.50
SD of Lane Position	0.25 (0.16)	0.29 (0.10)	1.73
Lane crossings per minute	0.39 (0.96)	0.85 (1.71)	2.40*

Performance of the proposed model at the mean values

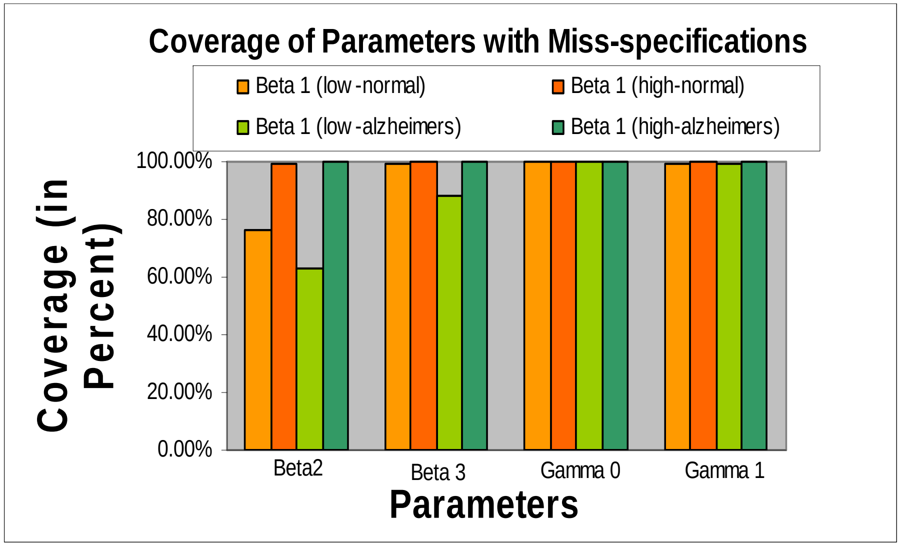
- We conducted a simulation study when we are under the true parameter setting, to see how well our model estimates the parameters. We report the coverage probabilities as a performance measuring stick (the higher the coverage the better).

Parameters $\beta_1 =$ $\beta_2 =$ $\beta_3 =$ $\gamma_0 =$ $\gamma_1 =$	Coverage probability
$\beta_1 = 0.05$ $\beta_1 = 0.052$ $\beta_2 = 0.47$ $\beta_2 = 0.31$ $\beta_3 = 0.48$ $\beta_3 = 0.64$ $\gamma_0 = 0.63$ $\gamma_0 = 0.42$ $\gamma_1 = 2.29$ $\gamma_1 = 1.63$	100% (100%) 97% (99%) 100% (100%) 100% (100%) 100% (100%)

- We explored the coverage probability when one parameter is miss-specified.
  - β_1 set at low level; repercussion on coverage probability
  - β_1 set at high level; repercussion on coverage probability
- We want to see how well re-centering parameter is covered.

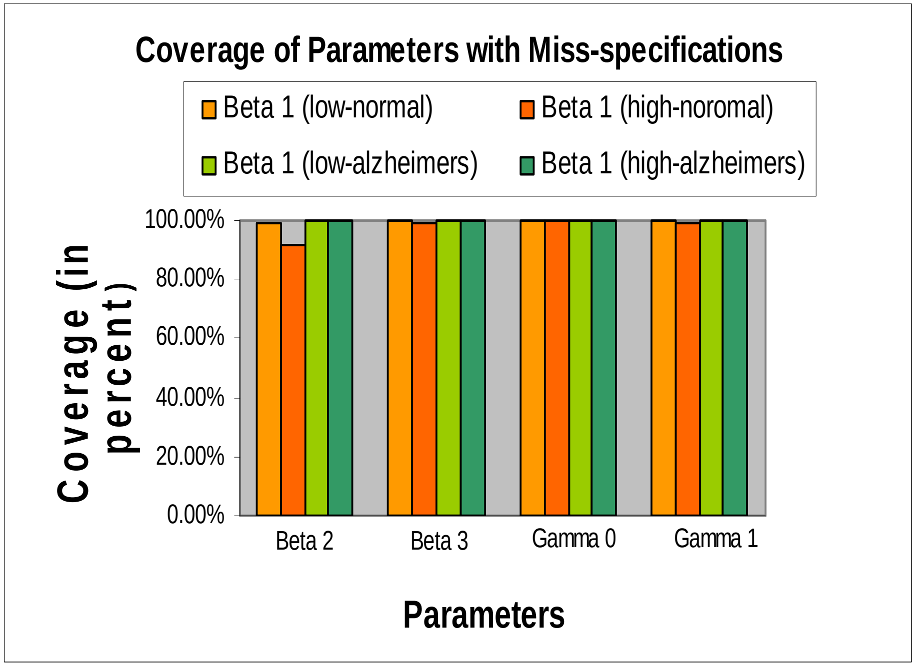
Performance of the proposed model under various parameter miss-specifications

Parameters miss-specification	$\beta_{1\ (Low)} = 0.036$ $\beta_{1\ (Low)} = 0.031$	$\beta_{1\ (high)} = 0.07$ $\beta_{1\ (high)} = 0.073$
β_2	76.5% (63%)	99.5% (100%)
β_3	92.5% (88.5%)	100% (100%)
γ_0	100% (100%)	100% (100%)
γ_1	99% (99.5%)	100% (100%)



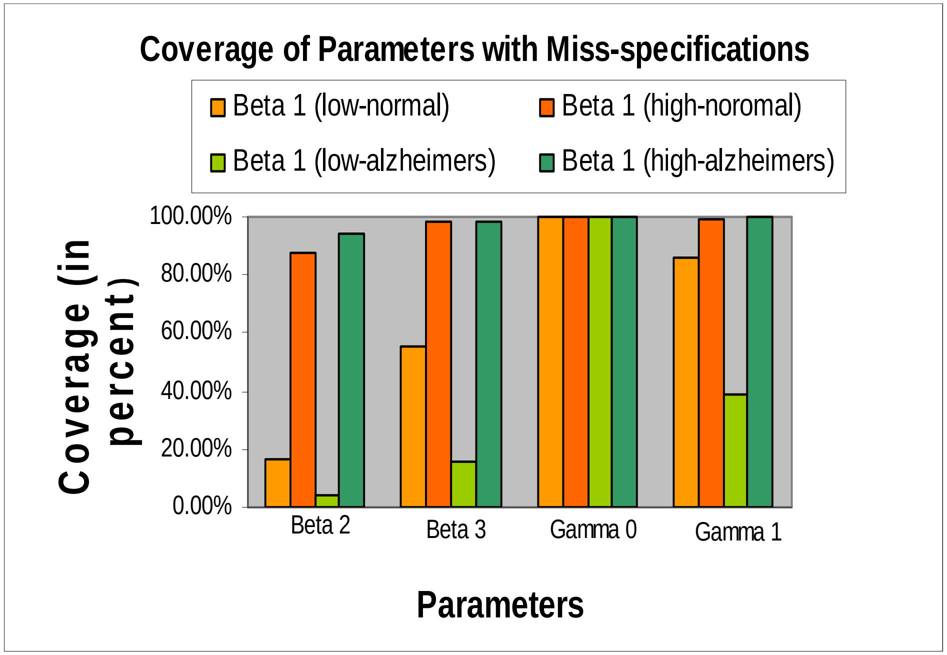
Performance of the proposed model under various parameter miss-specifications

Parameters miss-specification	$\beta_{1\,(Low)} = 0.036$ $\beta_{1\,(Low)} = 0.031$	$\beta_{1\,(high)} = 0.07$ $\beta_{1\,(high)} = 0.073$
β_2	99.5% (100%)	91.5% (100%)
β_3	100% (100%)	99% (100%)
$\gamma_{0\,(low)}$	100% (100%)	100% (100%)
$\gamma_{1\,(low)}$	100% (100%)	99.5% (100%)



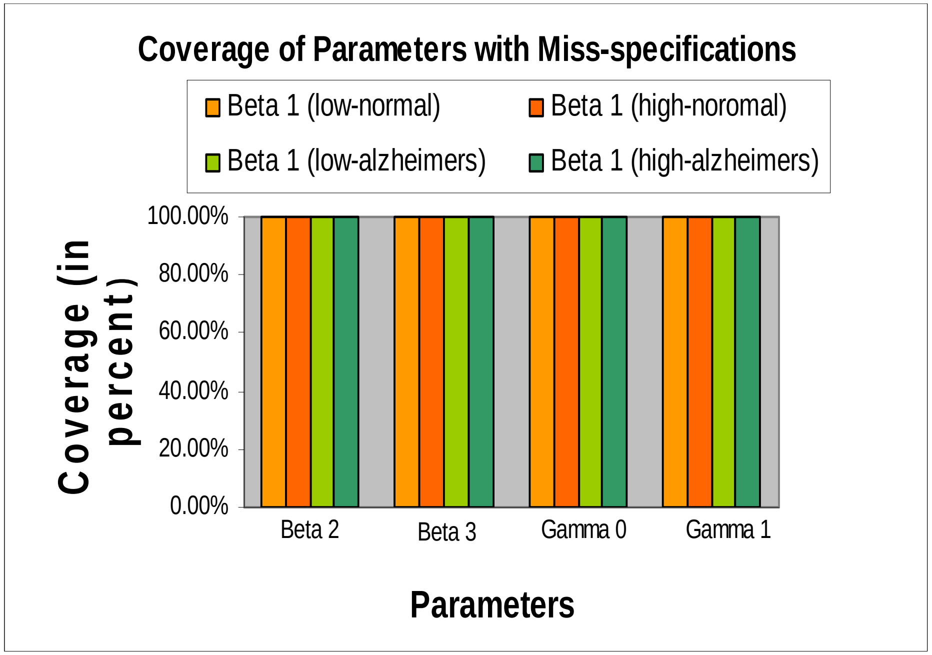
Performance of the proposed model under various parameter miss-specifications

Parameters miss-specification	$\beta_{1\ (Low)}=0.036$ $\beta_{1\ (Low)}=0.03$	$\beta_{1\ (high)}=0.07$ $\beta_{1\ (high)}=0.07$
β_2	16.5% (4%)	88% (94.5%)
β_3	55% (16%)	98% (98.5%)
$\gamma_{0\ (low)}$	100% (100%)	100% (100%)
$\gamma_{1\ (high)}$	86% (39%)	99% (100%)



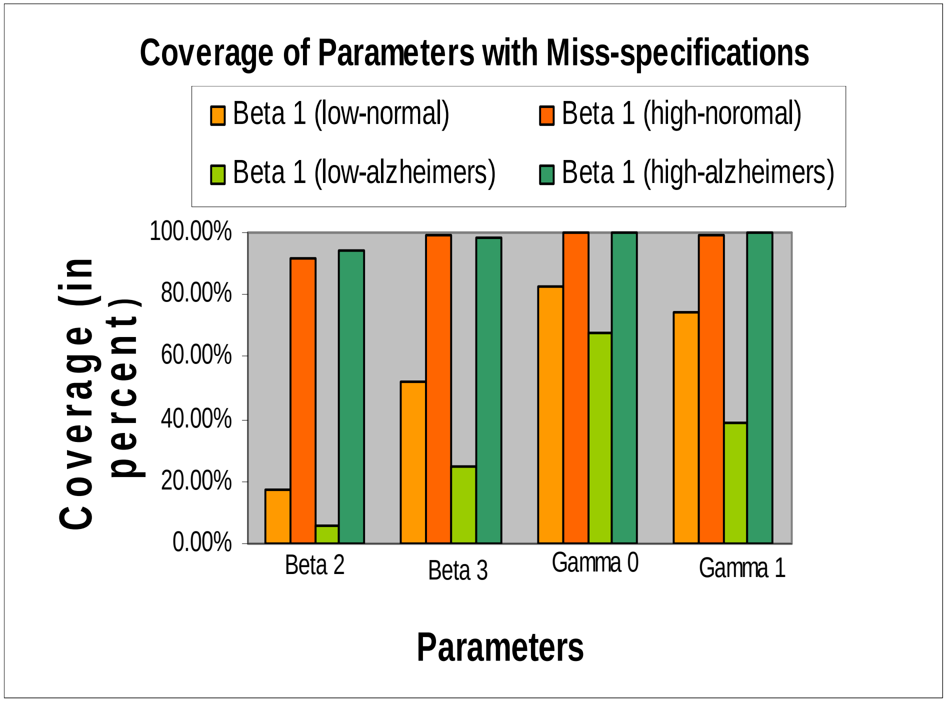
Performance of the proposed model under various parameter miss-specifications

Parameters miss-specification	$\beta_{1\ (Low)} = 0.036$ $\beta_{1\ (Low)} = 0.031$	$\beta_{1\ (high)} = 0.07$ $\beta_{1\ (high)} = 0.073$
β_2	100% (100%)	100% (100%)
β_3	100% (100%)	100% (100%)
$\gamma_{0\ (high)}$	100% (100%)	100% (100%)
$\gamma_{1\ (low)}$	100% (100%)	100% (100%)



Performance of the proposed model under various parameter miss-specifications

Parameters miss-specification	$\beta_{1\ (Low)} = 0.036$ $\beta_{1\ (Low)} = 0.031$	$\beta_{1\ (high)} = 0.07$ $\beta_{1\ (high)} = 0.073$
β_2	17.5% (6%)	91.5% (94%)
β_3	52% (25%)	99% (98%)
$\gamma_{0\ (high)}$	82.5% (68%)	100% (100%)
$\gamma_{1\ (high)}$	74.5% (39%)	99.5% (100%)



Conclusions

- Overall, the model covers the normal patients driving behaviors better than the Alzheimer's.
- For Alzheimer patients, we experience the worst performance when the flat component of the projection is set to one standard deviation below the average. This is to say that our model predicts that subjects with Alzheimer's will have more of a flat behavior than expected. That is to say if we were to construct a device to help Alzheimer's subjects, we will want to make sure the device puts more weight on the flat.
- As for re-centering, our model generally covers re-centering for neurologically normal subjects but not for Alzheimer's subjects.
- We might note that when poor coverage occurs in linear and quadratic projection, the re-centering parameters are poorly covered. To use our model to assist neurologically impaired individuals, we would have the device powered to recognize linear and quadratic projections that may have fallen below the average specification and advise the driver to re-center.

Future Work

- In the future, we will want to find the best setting by using a ROC curve.
- Instead of setting parameters high and low to see how much we cover them, we can use the mean and generate random variances around the mean corresponding to subject-specific variability and see how well the model performs.

Acknowledgements

- Dr. Gideon Zamba, Ph.D
- Dr. Jeffrey Dawson, Ph.D
- Biostatistics Department at the University of Iowa
- A.V.A.
- National Science Foundation