

Department of Statistics and Actuarial Sciences
University of Western Ontario
presented by Hossein

The Design of Simulation Studies in Medical Statistics

Biostatistics 9510 presentation

November 17, 2016



Introduction

Issues to consider when designing a simulation study

Review of some papers

DISCUSSION

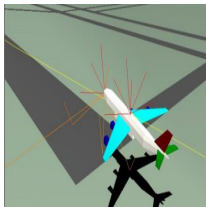
Why SIMULATION?



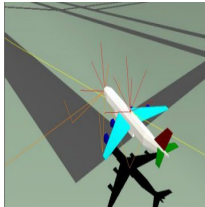
Why SIMULATION?



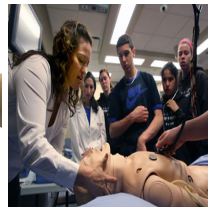
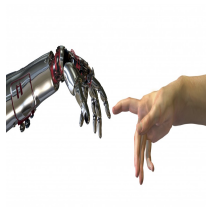
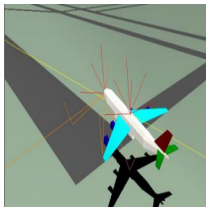
Why SIMULATION?



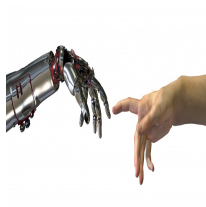
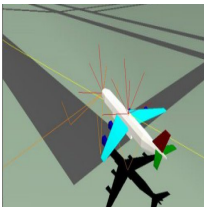
Why SIMULATION?



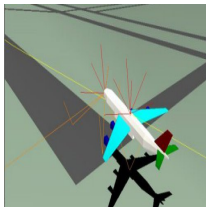
Why SIMULATION?



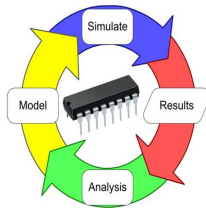
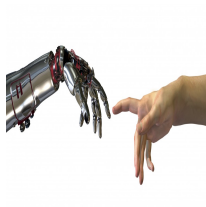
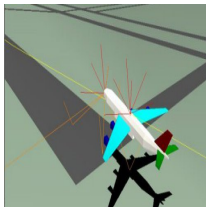
Why SIMULATION?



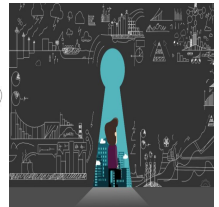
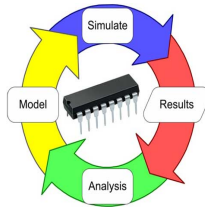
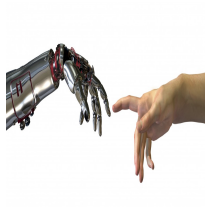
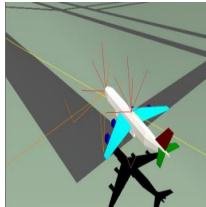
Why SIMULATION?



Why SIMULATION?



Why SIMULATION?





- ▶ Simulation studies use computer intensive procedures to assess the performance of a variety of statistical methods in relation to a known truth.



- ▶ Simulation studies use computer intensive procedures to assess the performance of a variety of statistical methods in relation to a known truth.
- ▶ High-quality simulations to reflect the complex situations (in practice)



- ▶ Simulation studies use computer intensive procedures to assess the performance of a variety of statistical methods in relation to a known truth.
- ▶ High-quality simulations to reflect the complex situations (in practice)
- ▶ Hard to achieve such results with real data



- ▶ Simulation studies use computer intensive procedures to assess the performance of a variety of statistical methods in relation to a known truth.
- ▶ High-quality simulations to reflect the complex situations (in practice)
- ▶ Hard to achieve such results with real data
- ▶ Know the criteria for designing a good quality **simulation study**:



- ▶ Simulation studies use computer intensive procedures to assess the performance of a variety of statistical methods in relation to a known truth.
- ▶ High-quality simulations to reflect the complex situations (in practice)
- ▶ Hard to achieve such results with real data
- ▶ Know the criteria for designing a good quality **simulation study**:
 - ▶ Designed so that performing many real studies



- ▶ Simulation studies use computer intensive procedures to assess the performance of a variety of statistical methods in relation to a known truth.
- ▶ High-quality simulations to reflect the complex situations (in practice)
- ▶ Hard to achieve such results with real data
- ▶ Know the criteria for designing a good quality **simulation study**:
 - ▶ Designed so that performing many real studies
 - ▶ Sufficient details provided



- ▶ Simulation studies use computer intensive procedures to assess the performance of a variety of statistical methods in relation to a known truth.
- ▶ High-quality simulations to reflect the complex situations (in practice)
- ▶ Hard to achieve such results with real data
- ▶ Know the criteria for designing a good quality **simulation study**:
 - ▶ Designed so that performing many real studies
 - ▶ Sufficient details provided
 - ▶ Involve careful consideration of all required aspects of the study

Issues to consider when designing a simulation study



- ▶ **Clearly defined aims and objectives**

Issues to consider when designing a simulation study



- ▶ **Clearly defined aims and objectives**
- ▶ **Simulation procedures**

Issues to consider when designing a simulation study



- ▶ **Clearly defined aims and objectives**
- ▶ **Simulation procedures**
 - ▶ Choice of random number generator

Issues to consider when designing a simulation study



- ▶ **Clearly defined aims and objectives**
- ▶ **Simulation procedures**
 - ▶ Choice of random number generator
 - ▶ Level of dependence between simulations

Issues to consider when designing a simulation study



- ▶ **Clearly defined aims and objectives**
- ▶ **Simulation procedures**
 - ▶ Choice of random number generator
 - ▶ Level of dependence between simulations
 - ▶ Allowance for failures

Issues to consider when designing a simulation study



- ▶ **Clearly defined aims and objectives**
- ▶ **Simulation procedures**
 - ▶ Choice of random number generator
 - ▶ Level of dependence between simulations
 - ▶ Allowance for failures
 - ▶ Starting seeds

Issues to consider when designing a simulation study



- ▶ **Clearly defined aims and objectives**
- ▶ **Simulation procedures**
 - ▶ Choice of random number generator
 - ▶ Level of dependence between simulations
 - ▶ Allowance for failures
 - ▶ Stratifying seeds
 - ▶ Software package

Issues to consider when designing a simulation study



- ▶ **Clearly defined aims and objectives**
- ▶ **Simulation procedures**
- ▶ **Methods for generating the data sets**

Issues to consider when designing a simulation study



- ▶ **Clearly defined aims and objectives**
- ▶ **Simulation procedures**
- ▶ **Methods for generating the data sets**
 - ▶ Univariate data

Issues to consider when designing a simulation study



- ▶ **Clearly defined aims and objectives**
- ▶ **Simulation procedures**
- ▶ **Methods for generating the data sets**
 - ▶ Univariate data
 - ▶ Multivariate data

Issues to consider when designing a simulation study



- ▶ **Clearly defined aims and objectives**
- ▶ **Simulation procedures**
- ▶ **Methods for generating the data sets**
 - ▶ Univariate data
 - ▶ Multivariate data
 - ▶ Time to event data

Issues to consider when designing a simulation study



- ▶ **Clearly defined aims and objectives**
- ▶ **Simulation procedures**
- ▶ **Methods for generating the data sets**

Issues to consider when designing a simulation study



- ▶ **Clearly defined aims and objectives**
- ▶ **Simulation procedures**
- ▶ **Methods for generating the data sets**

Issues to consider when designing a simulation study



- ▶ **Clearly defined aims and objectives**
- ▶ **Simulation procedures**
- ▶ **Methods for generating the data sets**

Issues to consider when designing a simulation study



- ▶ **Clearly defined aims and objectives**
- ▶ **Simulation procedures**
- ▶ **Methods for generating the data sets**
- ▶ **Scenarios to be investigated and methods for evaluation**

Issues to consider when designing a simulation study



- ▶ **Clearly defined aims and objectives**
- ▶ **Simulation procedures**
- ▶ **Methods for generating the data sets**
- ▶ **Scenarios to be investigated and methods for evaluation**
- ▶ **Estimates obtained from each simulation**

Issues to consider when designing a simulation study



- ▶ **Clearly defined aims and objectives**
- ▶ **Simulation procedures**
- ▶ **Methods for generating the data sets**
- ▶ **Scenarios to be investigated and methods for evaluation**
- ▶ **Estimates obtained from each simulation**
- ▶ **Number of simulations required**

Issues to consider when designing a simulation study



- ▶ Clearly defined aims and objectives
- ▶ Simulation procedures
- ▶ Methods for generating the data sets
- ▶ Scenarios to be investigated and methods for evaluation
- ▶ Estimates obtained from each simulation
- ▶ Number of simulations required
- ▶ Evaluating the performance of statistical methods for different scenarios

Issues to consider when designing a simulation study



- ▶ **Clearly defined aims and objectives**
- ▶ **Simulation procedures**
- ▶ **Methods for generating the data sets**
- ▶ **Scenarios to be investigated and methods for evaluation**
- ▶ **Estimates obtained from each simulation**
- ▶ **Number of simulations required**
- ▶ **Evaluating the performance of statistical methods for different scenarios**
 - ▶ **Assessment of bias**

Issues to consider when designing a simulation study



- ▶ **Clearly defined aims and objectives**
- ▶ **Simulation procedures**
- ▶ **Methods for generating the data sets**
- ▶ **Scenarios to be investigated and methods for evaluation**
- ▶ **Estimates obtained from each simulation**
- ▶ **Number of simulations required**
- ▶ **Evaluating the performance of statistical methods for different scenarios**
 - ▶ Assessment of bias
 - ▶ Assessment of accuracy

Issues to consider when designing a simulation study



- ▶ **Clearly defined aims and objectives**
- ▶ **Simulation procedures**
- ▶ **Methods for generating the data sets**
- ▶ **Scenarios to be investigated and methods for evaluation**
- ▶ **Estimates obtained from each simulation**
- ▶ **Number of simulations required**
- ▶ **Evaluating the performance of statistical methods for different scenarios**
 - ▶ Assessment of bias
 - ▶ Assessment of accuracy
 - ▶ Power, type I and II errors

Issues to consider when designing a simulation study



- ▶ **Clearly defined aims and objectives**
- ▶ **Simulation procedures**
- ▶ **Methods for generating the data sets**
- ▶ **Scenarios to be investigated and methods for evaluation**
- ▶ **Estimates obtained from each simulation**
- ▶ **Number of simulations required**
- ▶ **Evaluating the performance of statistical methods for different scenarios**
 - ▶ Assessment of bias
 - ▶ Assessment of accuracy
 - ▶ Power, type I and II errors
 - ▶ Assessment of coverage

Issues to consider when designing a simulation study



- ▶ **Clearly defined aims and objectives**
- ▶ **Simulation procedures**
- ▶ **Methods for generating the data sets**
- ▶ **Scenarios to be investigated and methods for evaluation**
- ▶ **Estimates obtained from each simulation**
- ▶ **Number of simulations required**
- ▶ **Evaluating the performance of statistical methods for different scenarios**
- ▶ **Presentation of the simulation results**



270 papers (Statistics in Medicine Journal - 2004)

Criteria	Frequency
Simulation studies	57 (21%)
Random generator and seeds (stated)	1
Statistical software	21
Number of simulations	100 - 100000
Justification for # simulations	6
Justification for data generation	23

An abstract graphic consisting of several flowing, curved lines in shades of light blue and white. The lines originate from the left side and sweep towards the right, creating a sense of motion and fluidity. Some lines are thicker and more prominent, while others are thinner and more delicate. The overall effect is reminiscent of a stylized wave or a dynamic, organic shape.

CONCLUSION?!!



- ▶ **Simulation studies are more accessible** (technology)



- ▶ **Simulation studies are more accessible** (technology)
- ▶ **Many correct decisions required for any simulation study**



- ▶ **Simulation studies are more accessible** (technology)
- ▶ **Many correct decisions required for any simulation study**
- ▶ **There is no single correct solution**



- ▶ **Simulation studies are more accessible** (technology)
- ▶ **Many correct decisions required for any simulation study**
- ▶ **There is no single correct solution**
- ▶ **Monte Carlo methods (MCMC, etc.) encompass any technique of statistical sampling**



- ▶ **Simulation studies are more accessible** (technology)
- ▶ **Many correct decisions required for any simulation study**
- ▶ **There is no single correct solution**
- ▶ **Monte Carlo methods (MCMC, etc.) encompass any technique of statistical sampling**
- ▶ **Important considerations when designing a simulation study**



- ▶ **Simulation studies are more accessible** (technology)
- ▶ **Many correct decisions required for any simulation study**
- ▶ **There is no single correct solution**
- ▶ **Monte Carlo methods (MCMC, etc.) encompass any technique of statistical sampling**
- ▶ **Important considerations when designing a simulation study**
 - ▶ identification of the estimate of interest



- ▶ **Simulation studies are more accessible** (technology)
- ▶ **Many correct decisions required for any simulation study**
- ▶ **There is no single correct solution**
- ▶ **Monte Carlo methods (MCMC, etc.) encompass any technique of statistical sampling**
- ▶ **Important considerations when designing a simulation study**
 - ▶ identification of the estimate of interest
 - ▶ appropriate analysis



- ▶ **Simulation studies are more accessible** (technology)
- ▶ **Many correct decisions required for any simulation study**
- ▶ **There is no single correct solution**
- ▶ **Monte Carlo methods (MCMC, etc.) encompass any technique of statistical sampling**
- ▶ **Important considerations when designing a simulation study**
 - ▶ identification of the estimate of interest
 - ▶ appropriate analysis
 - ▶ methods for comparison



- ▶ **Simulation studies are more accessible** (technology)
- ▶ **Many correct decisions required for any simulation study**
- ▶ **There is no single correct solution**
- ▶ **Monte Carlo methods (MCMC, etc.) encompass any technique of statistical sampling**
- ▶ **Important considerations when designing a simulation study**
 - ▶ identification of the estimate of interest
 - ▶ appropriate analysis
 - ▶ methods for comparison
 - ▶ criteria for evaluation these methods



- ▶ **Simulation studies are more accessible** (technology)
- ▶ **Many correct decisions required for any simulation study**
- ▶ **There is no single correct solution**
- ▶ **Monte Carlo methods (MCMC, etc.) encompass any technique of statistical sampling**
- ▶ **Important considerations when designing a simulation study**
 - ▶ identification of the estimate of interest
 - ▶ appropriate analysis
 - ▶ methods for comparison
 - ▶ criteria for evaluation these methods
 - ▶ number of situations to consider



- ▶ **Simulation studies are more accessible** (technology)
- ▶ **Many correct decisions required for any simulation study**
- ▶ **There is no single correct solution**
- ▶ **Monte Carlo methods (MCMC, etc.) encompass any technique of statistical sampling**
- ▶ **Important considerations when designing a simulation study**
 - ▶ identification of the estimate of interest
 - ▶ appropriate analysis
 - ▶ methods for comparison
 - ▶ criteria for evaluation these methods
 - ▶ number of situations to consider
 - ▶ reporting the results



THANK YOU!