

# Discrete Event Simulation in R to Support Healthcare Decision Making

Welcome & Course Introduction

**H. Koffijberg, PhD, MSc**

University of Twente  
Enschede, The Netherlands

**K. Degeling, PhD, MSc**

University of Melbourne  
Melbourne, Australia

**M. van de Ven, MSc**

OPEN Health  
Rotterdam, The Netherlands

# Getting started!

- Introduction of the lecturers
- Course objectives
- Course outline

DISCRETE EVENT SIMULATION – SMDM 2021 SHORT COURSE  
H. KOFFIJBERG, K. DEGELING, M. VAN DE VEN

# Course lecturers



Erik Koffijberg



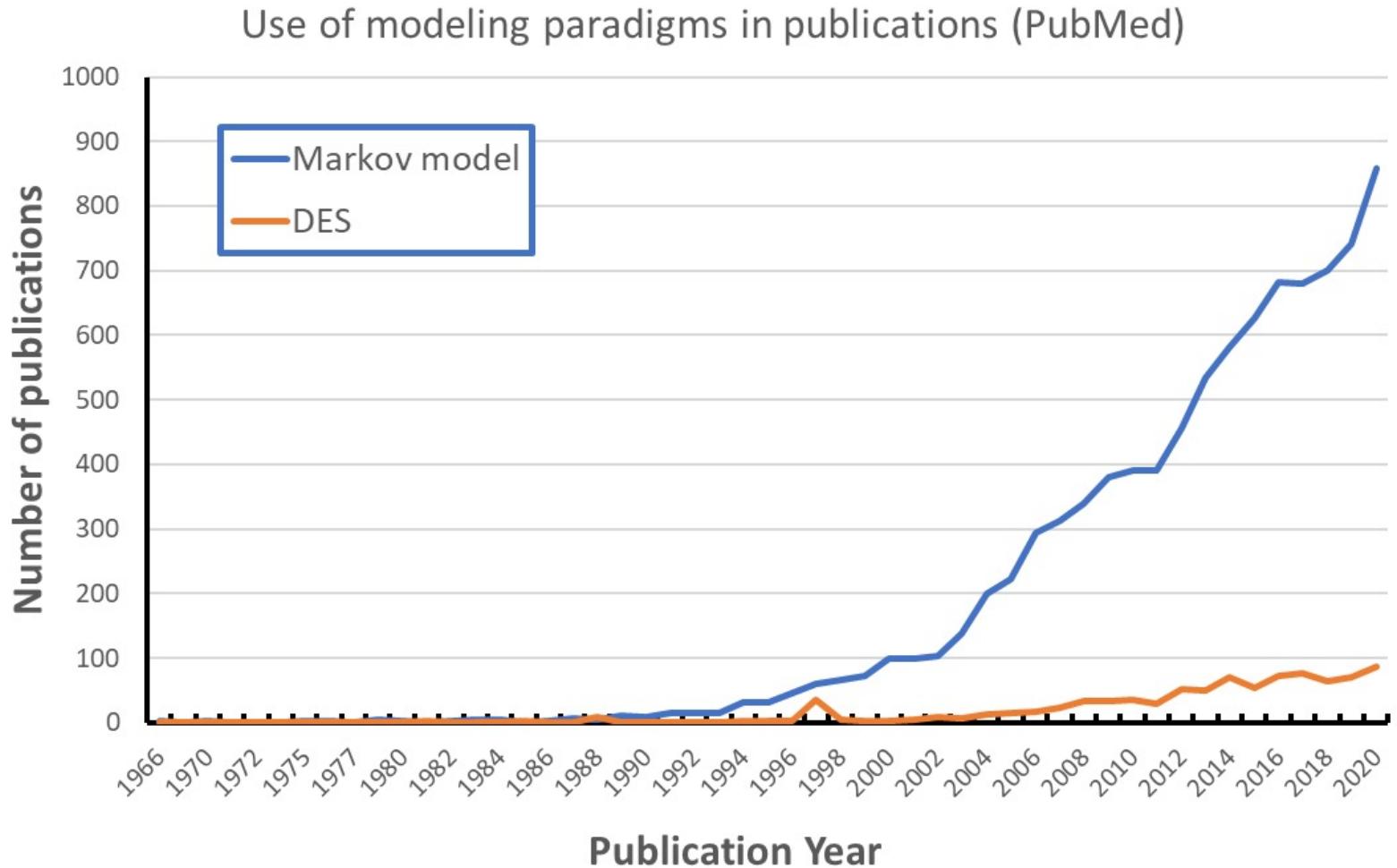
Koen Degeling



Michiel van de Ven

# Course background

- Modeling is becoming ever more important
- PubMed [tiab]: “Markov model” vs “Discrete Event Simulation”
- Modeling paradigm mostly determined by
  - Experience of modeller
  - Ease of use
  - Computing power
  - Availability of data
  - Complexity of (clinical) practice



# Course background

- Last decade and future is characterized by
  - Increasing computational power
  - Increasing availability of (highly detailed) patient-level data
  - Increasing complexity of clinical practice
- This makes it **more feasible and necessary** to apply flexible modeling paradigms to create realistic evidence-based models to support decision making in healthcare
- DES is one of the most flexible modeling paradigms available
  - The use of DES will increase *if* modelers gain first hand experience with DES, and experience its advantages over more conventional modeling paradigms

# Course objectives

*After the completion of this course, participants will be able to:*

1. Understand the design choices necessary for developing DES.
2. Recognize the different types of evidence to populate DES.
3. Distinguish between alternative approaches for implementing competing events in DES.
4. Distinguish between different types of uncertainty in DES.
5. Understand how a basic DES can be implemented in R using the `simmer` package.

# Course outline

- Presentation 1 - Introduction to DES
  - Demonstration 1 – Using simmer for DES in R
  - Presentation 2
    - *Modeling competing events*
    - *Patient-level versus aggregate data*
- 
- Presentation 3 – Uncertainty in DES
  - Demonstration 2 – Implementing probabilistic analysis
  - Discussion with Q&A





120-1 SHORT COURSE  
M. VAN DE VEN



# Discrete Event Simulation in R to Support Healthcare Decision Making

The basics of discrete event simulation

**H. Koffijberg, PhD, MSc**

University of Twente  
Enschede, The Netherlands

**K. Degeling, PhD, MSc**

University of Melbourne  
Melbourne, Australia

**M. van de Ven, MSc**

OPEN Health  
Rotterdam, The Netherlands

# Outline

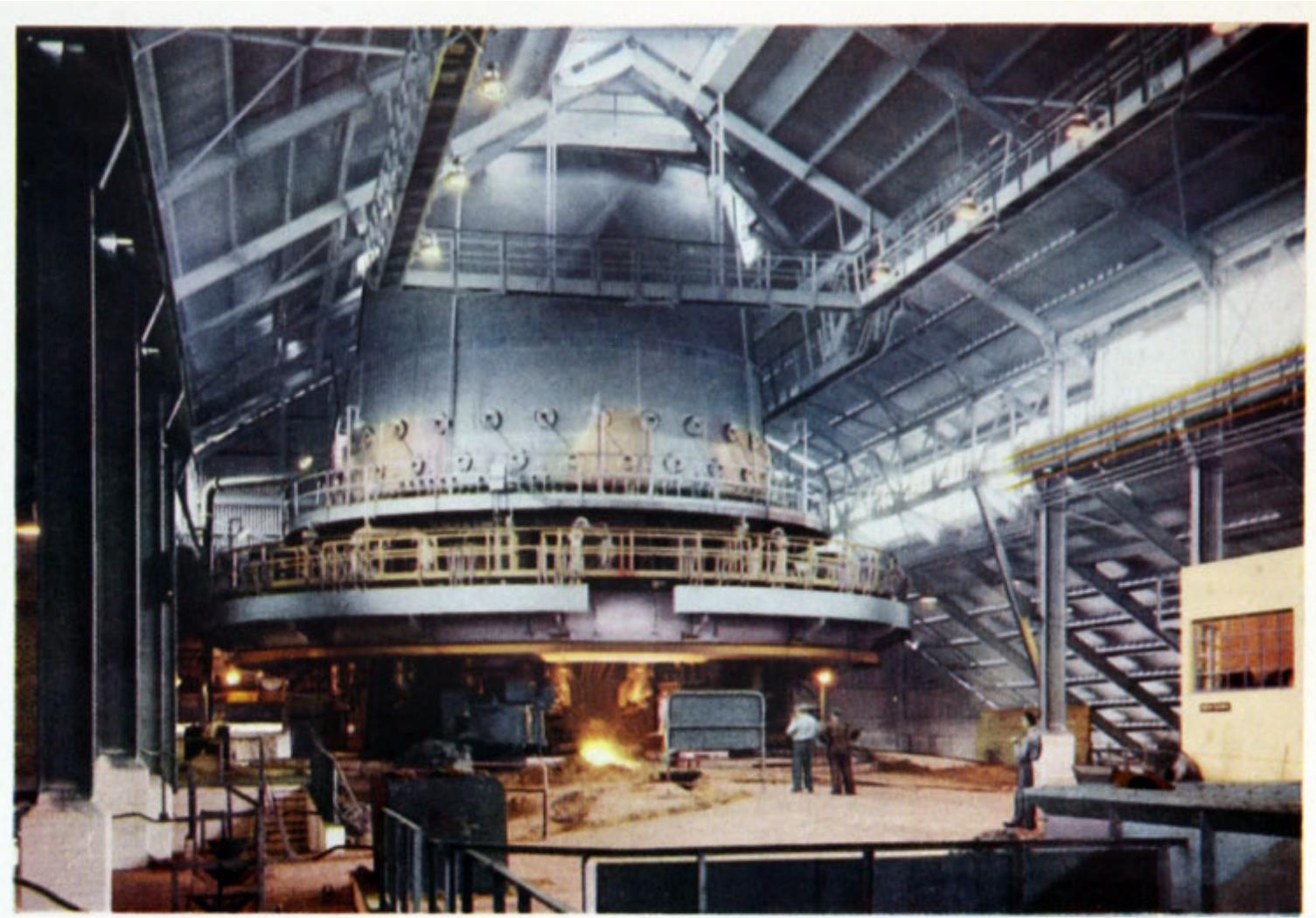
1. The basics of time-to-event-modeling
2. DES compared to other model types
3. An overview of potential applications in healthcare

DISCRETE EVENT SIMULATION – SMDA 2021 SHORT COURSE  
H. KOFFIJBERG, K. DEGELING, M. VAN DE VEN

# **1. The basics of time-to-event-modeling**

DISCRETE EVENT SIMULATION – SMIDM 2021 SHORT COURSE  
H. KOFFIJBERG, K. DEGELING, M. VAN DE VEN

# Discrete Event Simulation



[https://www.gracesguide.co.uk/United\\_Steel\\_Companies](https://www.gracesguide.co.uk/United_Steel_Companies)

# Discrete event simulation (DES)

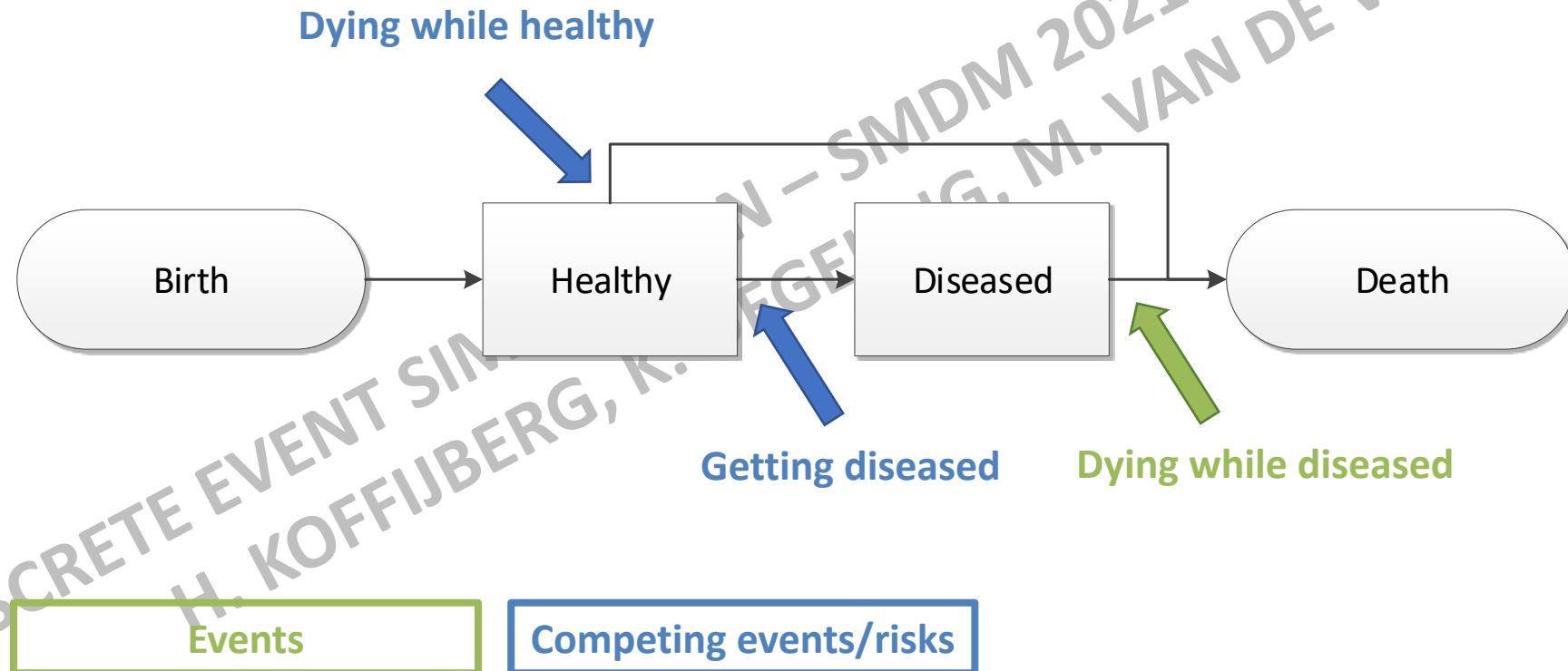
- Modelling of operations systems
- A simulation method used to characterize and analyze specific processes and use of resources
- Most problems involve:
  - Resource utilization
  - Queues, i.e., waiting times
- DES can be useful to also analyze effects on health-related outcomes

# Discrete event simulation (DES)

- From the behavior of the system to an ordered sequence of events
- Entities can experience events at any discrete time after the previous event
- Events can occur at any time
- Extensive reading: *Law A.M., Simulation Modeling and Analysis, McGraw-Hill Higher Education, 2014.*

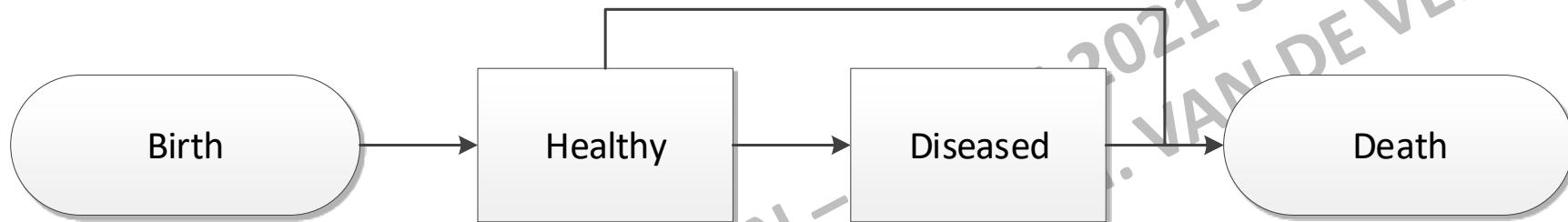
# Time to event modeling

- Basic DES structure:

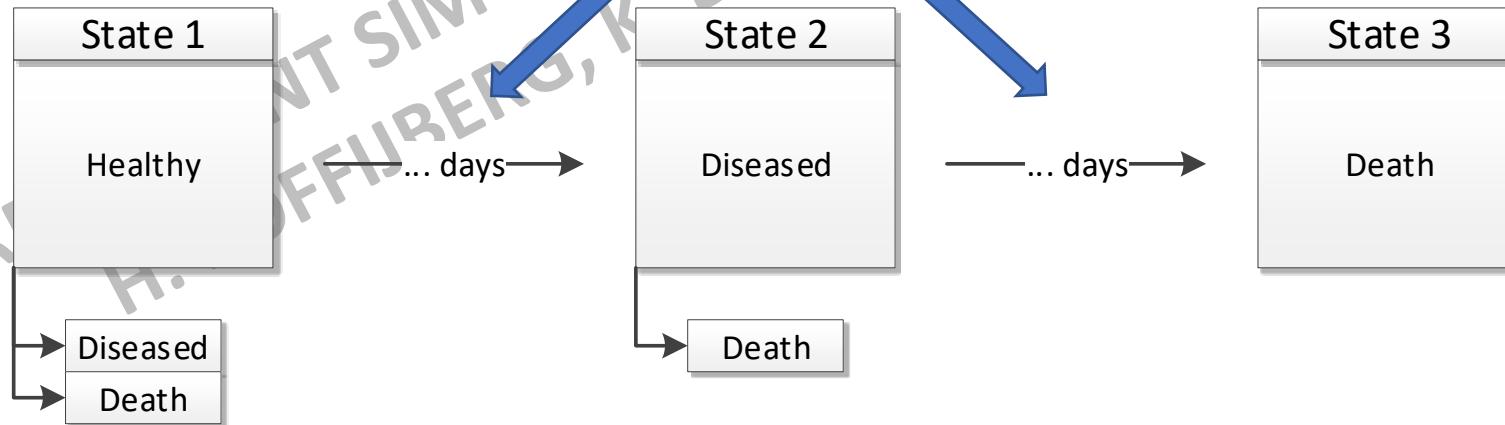


# Time to event modeling

- Basic DES structure & simulation flow:



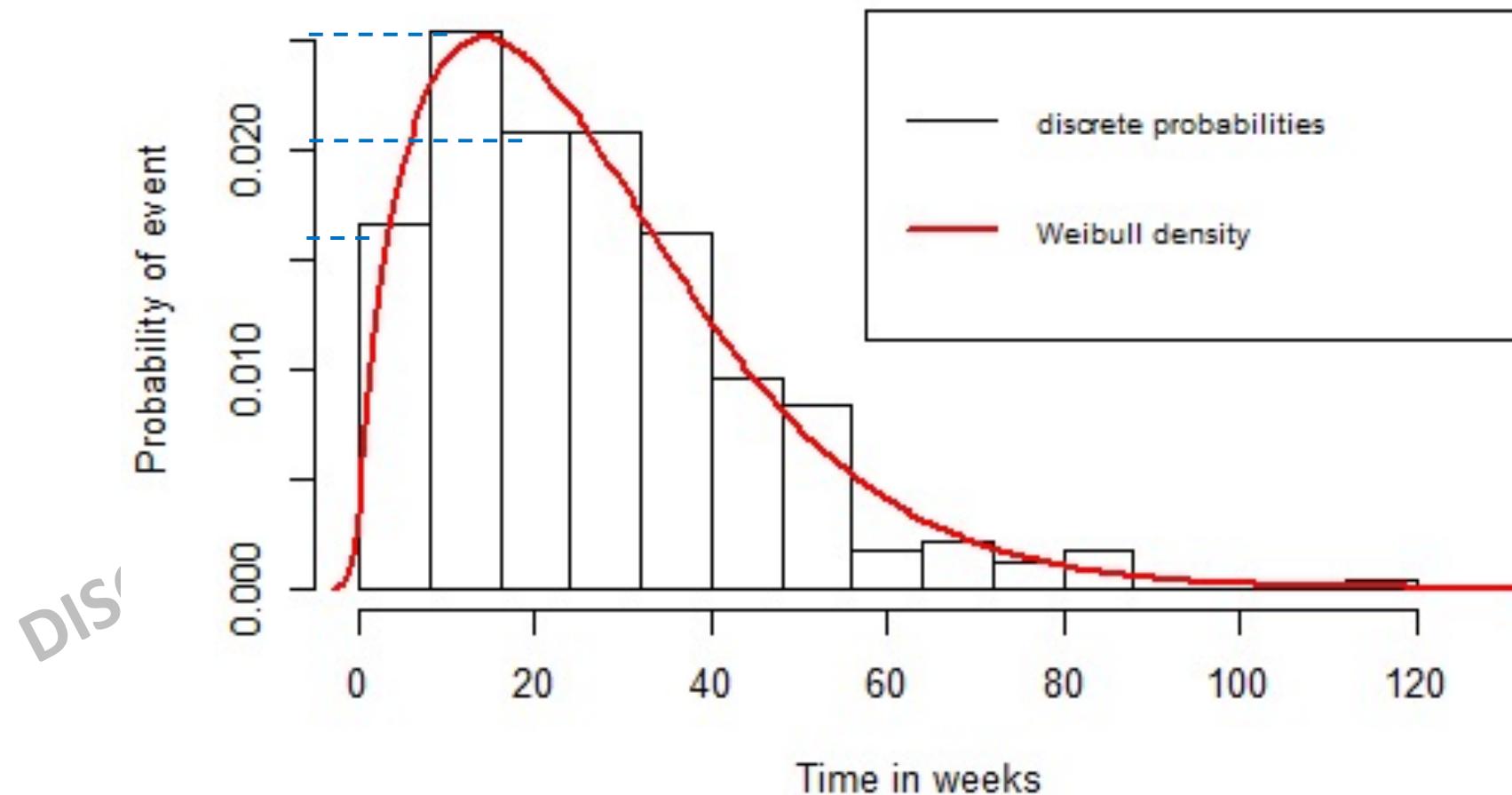
Example:



# Time to event modeling vs discrete time cycles

Illustration with hypothetical data on time to disease progression:

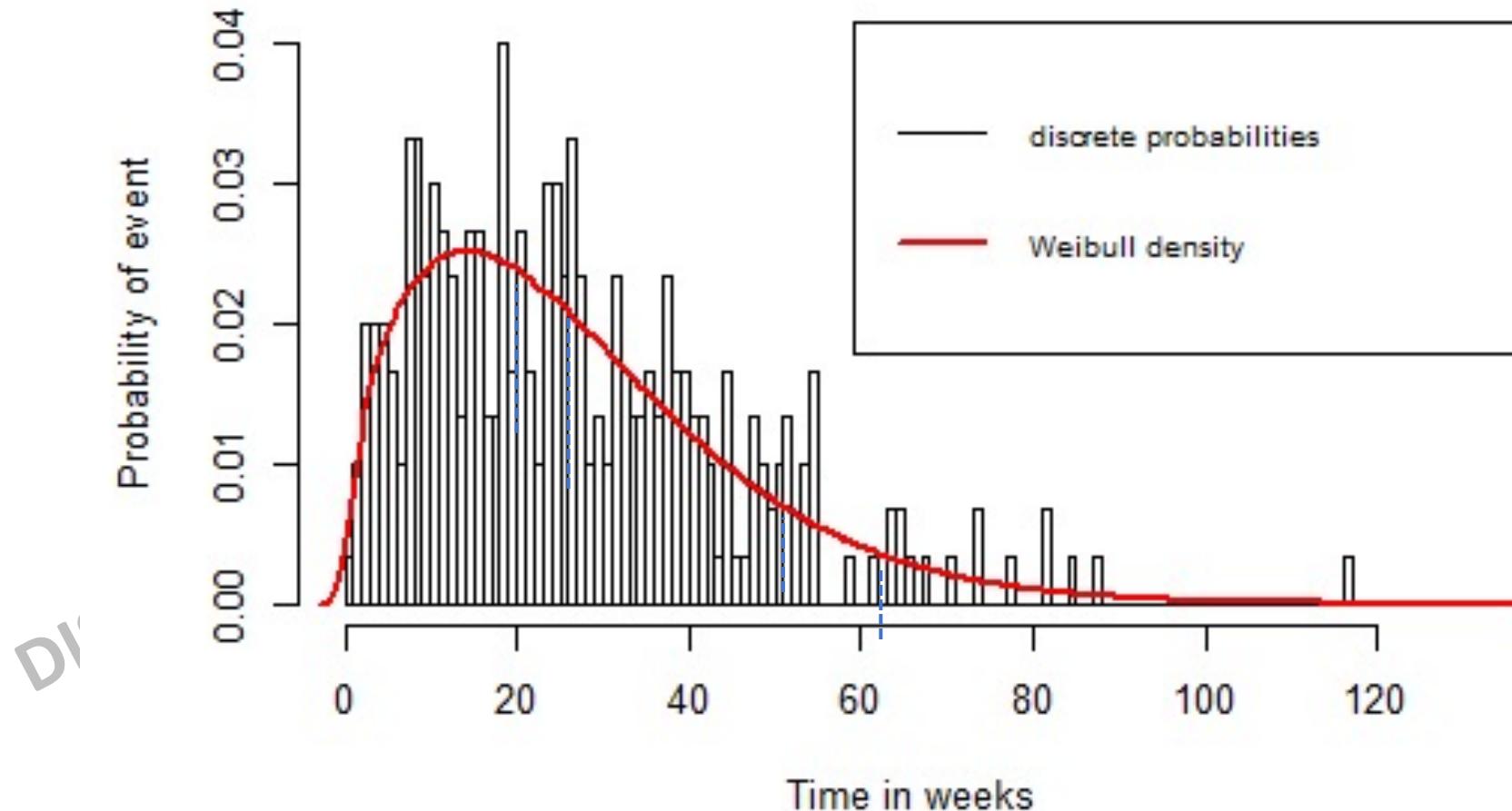
- *Observations for 300 individuals, no missings, cycle size = 8 weeks*



# Time to event modeling vs discrete time cycles

Illustration with hypothetical data on time to disease progression:

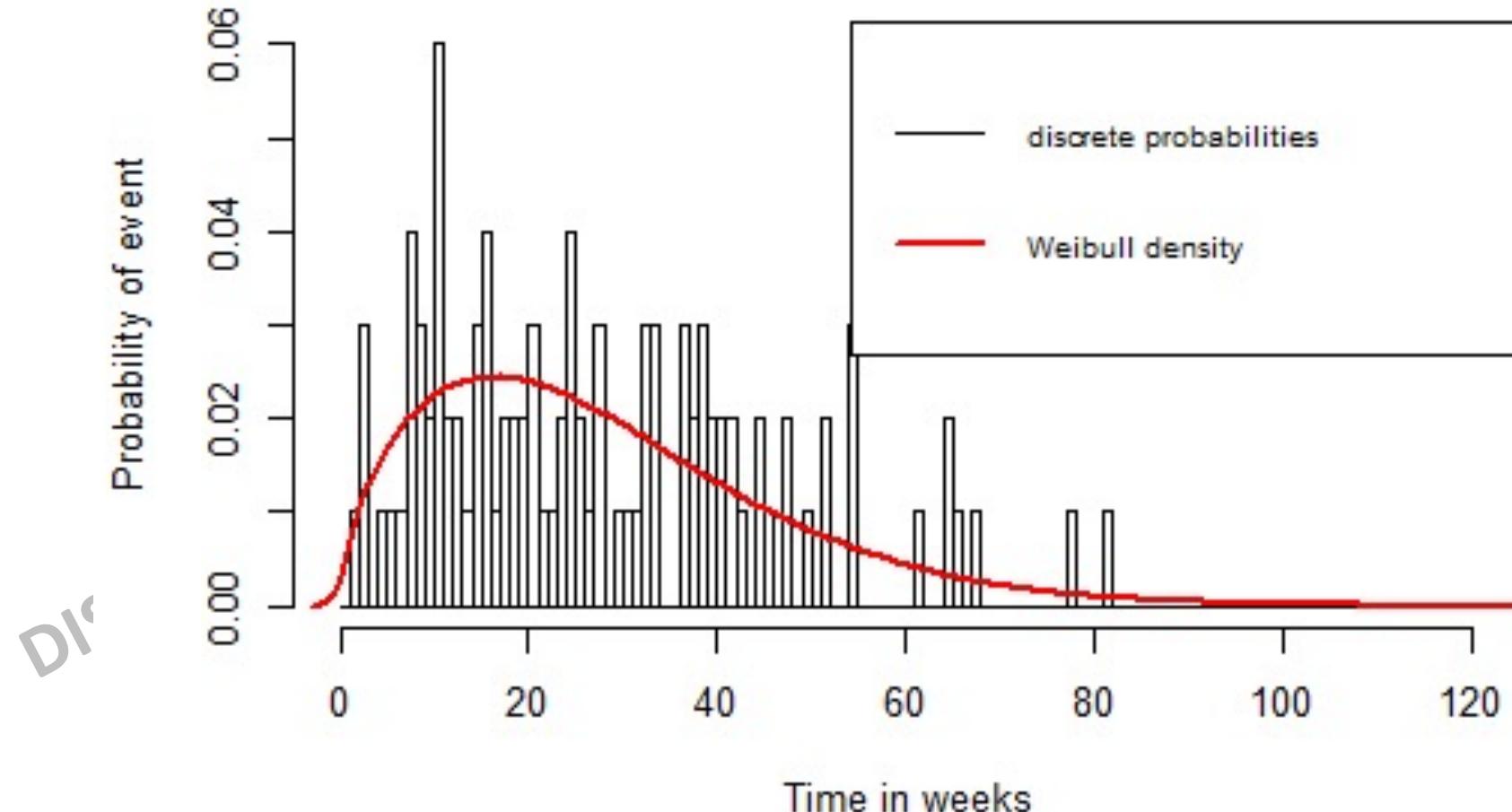
- *Observations for 300 individuals, no missings, cycle size = 1 week*



# Time to event modeling vs discrete time cycles

Illustration with hypothetical data on time to disease progression:

- *Observations for 100 individuals, no missings, cycle size = 1 week*



# Time to event modeling vs discrete time cycles

- Empirical time (to-event) data can be directly used
  - Discretization unnecessary – unless required by the model structure
- Discretization for PL-STM is possible but *undesirable*
- Discretization also feasible (theoretically) using a table populated with values from distribution

## **2. DES compared to other model types**

Discrete time cycles vs. continuous time, and other dynamic simulation model types

DISCRETE EVENT SIMULATION - ISNDM 2021 SHORT COURSE  
H. KOFFIJBERG, K. DEGELING, M. VAN DE VEN

# DES compared with cohort Markov model (DT-STM)

COURSE

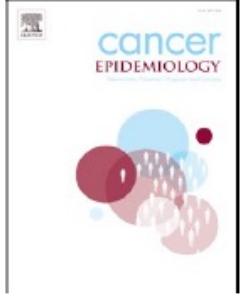
Cancer Epidemiology 57 (2018) 60–67

Contents lists available at ScienceDirect

Cancer Epidemiology

journal homepage: [www.elsevier.com/locate/canep](http://www.elsevier.com/locate/canep)

ELSEVIER



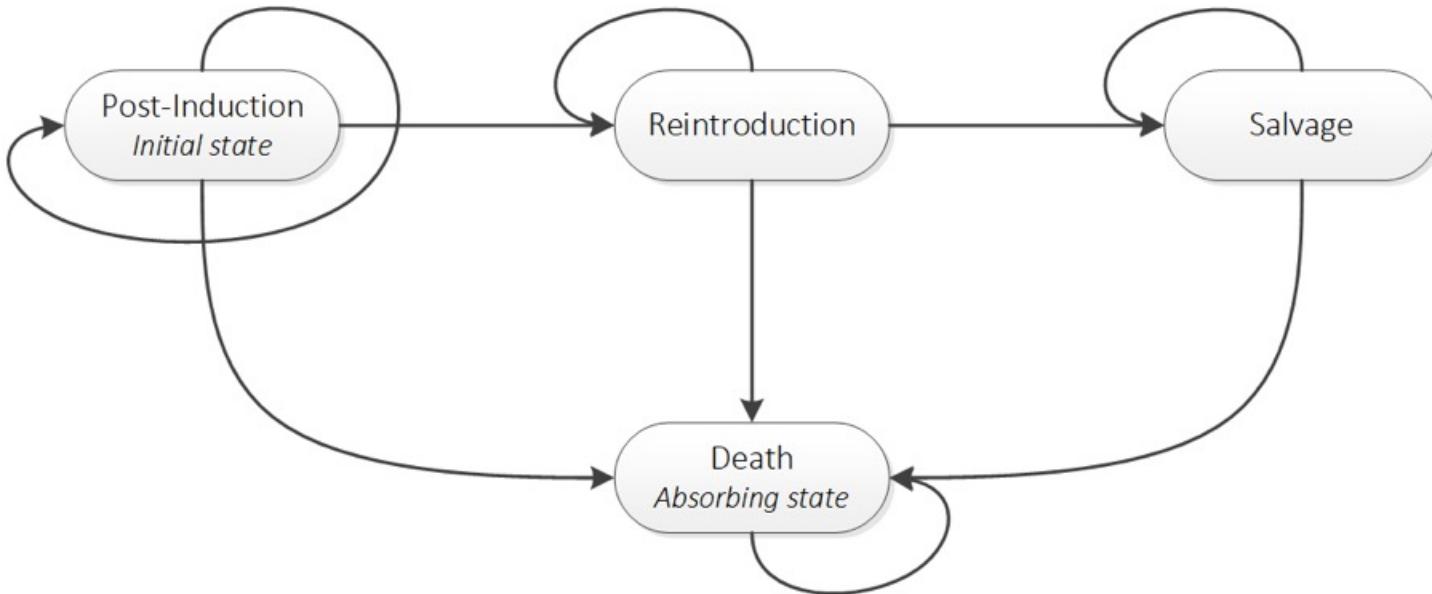
Check for updates

Matching the model with the evidence: comparing discrete event simulation and state-transition modeling for time-to-event predictions in a cost-effectiveness analysis of treatment in metastatic colorectal cancer patients

Koen Degeling<sup>a,1</sup>, Mira D. Franken<sup>b,1</sup>, Anne M. May<sup>c</sup>, Martijn G.H. van Oijen<sup>d</sup>, Miriam Koopman<sup>b</sup>, Cornelis J.A. Punt<sup>d</sup>, Maarten J. IJzerman<sup>a</sup>, Hendrik Koffijberg<sup>a,c,\*</sup>

# DT-STM

a.



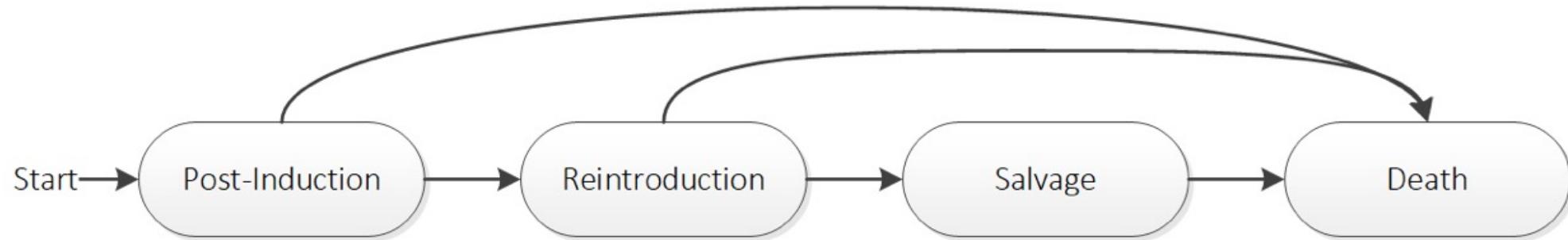
RSE

Example:

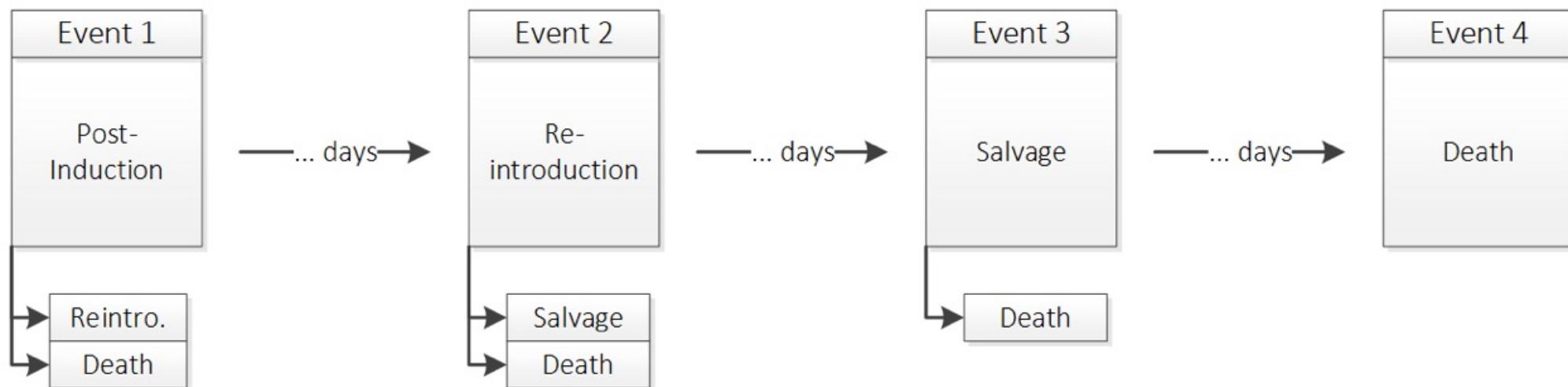
Cycle 1	Cycle 2	Cycle 3	Cycle 4	Cycle 5	Cycle 6	Cycle 7	... .. ..	Cycle N
Post- Induction	Post- Induction	Post- Induction	Re- introduction	Salvage	Salvage	Death		Death
<div style="display: flex; align-items: center;"><span style="border: 1px solid black; padding: 2px;">Post-Ind.</span><span style="margin: 0 10px;">→</span><span style="border: 1px solid black; padding: 2px;">Post-Ind.</span><span style="margin: 0 10px;">→</span><span style="border: 1px solid black; padding: 2px;">Post-Ind.</span><span style="margin: 0 10px;">→</span><span style="border: 1px solid black; padding: 2px;">Reintro.</span><span style="margin: 0 10px;">→</span><span style="border: 1px solid black; padding: 2px;">Salvage</span><span style="margin: 0 10px;">→</span><span style="border: 1px solid black; padding: 2px;">Death</span></div>	<div style="display: flex; align-items: center;"><span style="border: 1px solid black; padding: 2px;">Post-Ind.</span><span style="margin: 0 10px;">→</span><span style="border: 1px solid black; padding: 2px;">Post-Ind.</span><span style="margin: 0 10px;">→</span><span style="border: 1px solid black; padding: 2px;">Post-Ind.</span><span style="margin: 0 10px;">→</span><span style="border: 1px solid black; padding: 2px;">Reintro.</span><span style="margin: 0 10px;">→</span><span style="border: 1px solid black; padding: 2px;">Death</span></div>	<div style="display: flex; align-items: center;"><span style="border: 1px solid black; padding: 2px;">Post-Ind.</span><span style="margin: 0 10px;">→</span><span style="border: 1px solid black; padding: 2px;">Post-Ind.</span><span style="margin: 0 10px;">→</span><span style="border: 1px solid black; padding: 2px;">Post-Ind.</span><span style="margin: 0 10px;">→</span><span style="border: 1px solid black; padding: 2px;">Reintro.</span><span style="margin: 0 10px;">→</span><span style="border: 1px solid black; padding: 2px;">Death</span></div>	<div style="display: flex; align-items: center;"><span style="border: 1px solid black; padding: 2px;">Reintro.</span><span style="margin: 0 10px;">→</span><span style="border: 1px solid black; padding: 2px;">Salvage</span><span style="margin: 0 10px;">→</span><span style="border: 1px solid black; padding: 2px;">Death</span></div>	<div style="display: flex; align-items: center;"><span style="border: 1px solid black; padding: 2px;">Salvage</span><span style="margin: 0 10px;">→</span><span style="border: 1px solid black; padding: 2px;">Death</span></div>	<div style="display: flex; align-items: center;"><span style="border: 1px solid black; padding: 2px;">Salvage</span><span style="margin: 0 10px;">→</span><span style="border: 1px solid black; padding: 2px;">Death</span></div>	<div style="display: flex; align-items: center;"><span style="border: 1px solid black; padding: 2px;">Death</span></div>		<div style="display: flex; align-items: center;"><span style="border: 1px solid black; padding: 2px;">Death</span></div>

# DES

DISIRSE

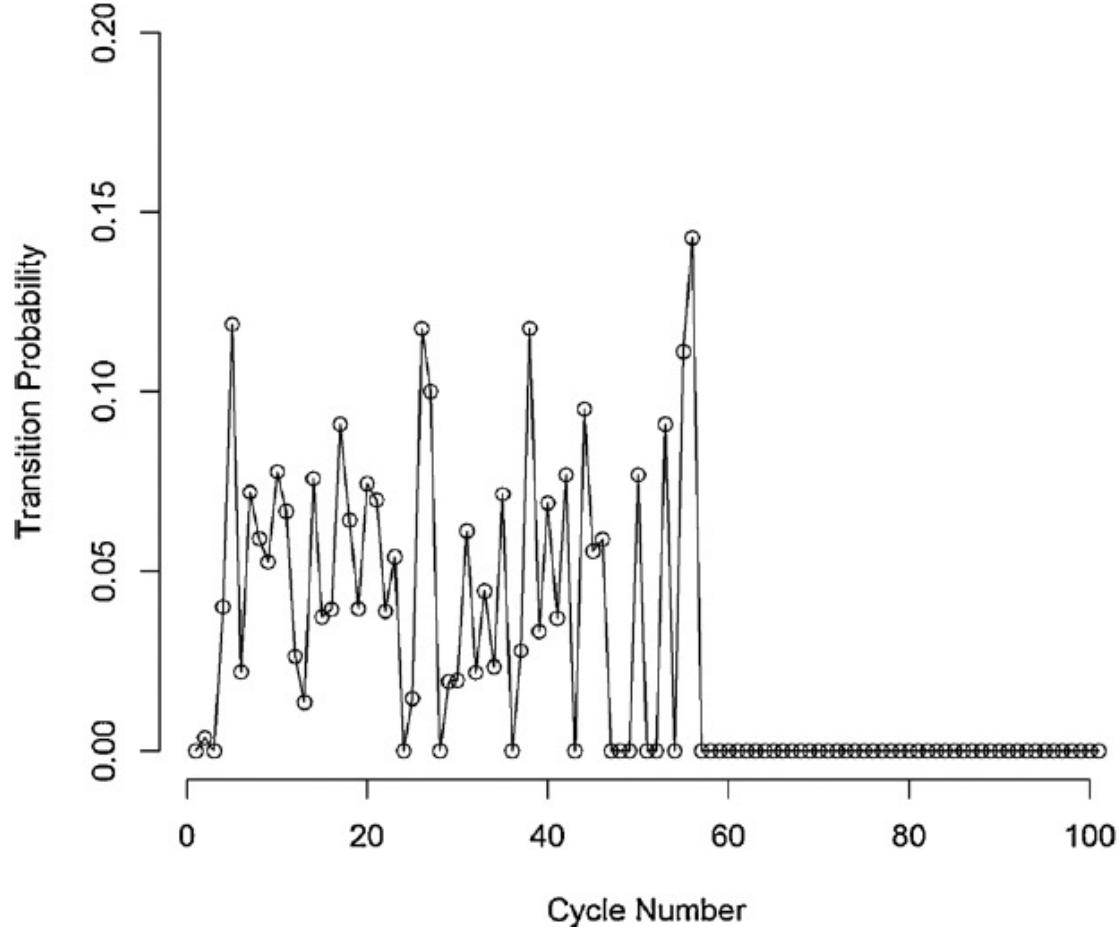


Example:



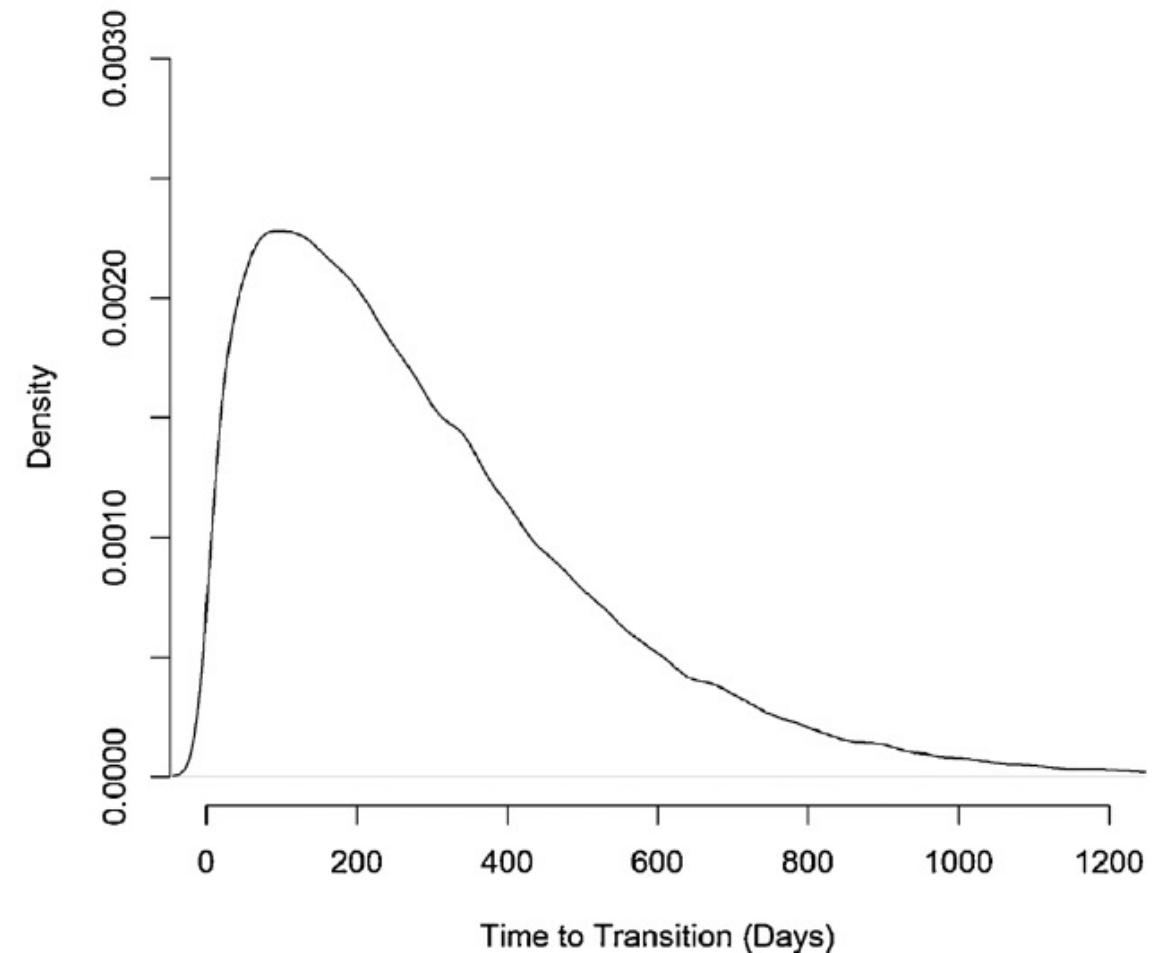
# Transitions for the maintenance strategy

a) Evidence Structure for the STM



**DT-STM:** probability curve for the time to transition of the post-introduction to the reintroduction health state per cycle (with a 3-week duration)

b) Evidence Structure for the DES model



**DES:** probability density curve for the time to transition of the post-introduction to the reintroduction health state

# DES compared with cohort Markov model (DT-STM)

## Results

- DT-STM: transition probabilities were irregular and sensitive to single events
- DES: parametric distributions resulted in smooth time-to-event curves.
- The DES model represented the trial data more accurately

# DES compared with DT-STM - Conclusion

- In this case study, the DT-STM and DES model did not yield substantially different outcomes
- However, when few events are observed per time cycle, DES is expected to yield a more accurate ICER.
- Which modeling method should be applied, in general, depends on the complexity of the clinical process to be modeled, the available evidence, and the modelers' experience.
- DES is preferable when patient-level time-to-event data is available

# DES compared to other simulation model types

- Other simulation model types that are applied in health economics are:
  - System Dynamics
  - Agent-based modeling
- System dynamics:
  - High level (i.e., aggregate-/population-level) modeling method
  - Typically used to understand the behavior of a complex system
- Agent-based modeling:
  - Model 'micro behavior' to understand 'macro effects'
  - Agent-based models consist of dynamically interacting rule-based agents

# Comparison of methods

Aspect	Markov	PSA	PL-STM	SD	DES	ABM
Entity Level	Cohort	Cohort	<u>Individual</u>	Cohort	<u>Individual</u>	<u>Individual</u>
Entity Role	Passive	Passive	Passive	Passive	Passive	<u>Active</u>
Population	Closed	Closed	Closed	Closed	<u>Open/Closed</u>	<u>Open/Closed</u>
Abstraction	Avg. Entity	Avg. Entity	Entity	System	Process	Entity
Structure	Health States	Health States	Health States	<u>System States</u>	<u>Process States</u>	Health States
Time Handling	Time Driven	Time Driven	Time Driven	Time Driven	<u>Event Driven</u>	Time Driven

- Also see “Marshall DA, et al. ISPOR Emerging Good Practices Task Force. Selecting a dynamic simulation modeling method for health care delivery research – Part 2. Value in Health. March 2015 18(2): 147-160” for a detailed comparison of SD, DES and ABM.

### **3. An overview of potential applications in healthcare**

DISCRETE EVENT SIMULATION – SMDM 2021 SHORT COURSE  
H. KOFFIJBERG, K. DEGELING, M. VAN DER VEN

# Applications in healthcare

- DES takes an operational/tactical perspective
- This is particularly relevant for:
  - System configurations
  - Resource planning
  - Scheduling
- Cost-effectiveness modeling studies

DISCRETE EVENT SIMULATION – SMDM 2021 SHORT COURSE  
H. KOFFELBERG, K. DEGELING, M. VAN DE VEN

# Applications in healthcare

- Example: Using simulation to determine the need for ICU beds for surgery patients

Philip Marc Troy, PhD, and Lawrence Rosenberg, MD, PhD, Montreal, Quebec, Canada

**Background.** As the need for surgical ICU beds at the hospital increases, the mismatch between demand and supply for those beds has led to the need to understand the drivers of ICU performance.

**Method.** A Monte Carlo simulation study of ICU performance was performed using a discrete event model that captured the events, timing, and logic of ICU patient arrivals and bed stays.

**Results.** The study found that functional ICU capacity, ie, the number of occupied ICU beds at which operative procedures were canceled if they were known to require an ICU stay, was the main determinant of the wait, the number performed, and the number of cancellations of operative procedures known to require an ICU stay. The study also found that actual and functional ICU capacity jointly explained ICU utilization and the mean number of patients that should have been in the ICU that were parked elsewhere.

**Conclusion.** The study demonstrated the necessity of considering actual and functional ICU capacity when analyzing surgical ICU bed requirements, and suggested the need for additional research on synchronizing demand with supply. The study also reinforced the authors' sense that simulation facilitates the evaluation of trade-offs between surgical management alternatives proposed by experts and the identification of unexpected drawbacks or opportunities of those proposals. (Surgery 2009;146:608-20.)

From The Sir Mortimer B. Davis Jewish General Hospital, Montreal, Quebec, Canada

Troy PM, Rosenberg L. Using simulation to determine the need for ICU beds for surgery patients. Surgery. 2009;146(4):608-17; discussion 17-20.

# Applications in healthcare

- Example (continued):
  - Improving ICU performance through:
    - Increasing the number of ICU beds for surgery patients
    - Improved alignment of the supply and demand of ICU beds
  - It involves efficient planning of scarce hospital resources to optimize waiting times
  - ICU bed capacity is requested from two sources:
    - Emergency department: Acute and elective surgery
    - Medical wards: Elective surgery and critical illness

# Applications in healthcare

ACE

- Example (continued):

DISCRETE EVENT S'  
H. KOFFIJ'

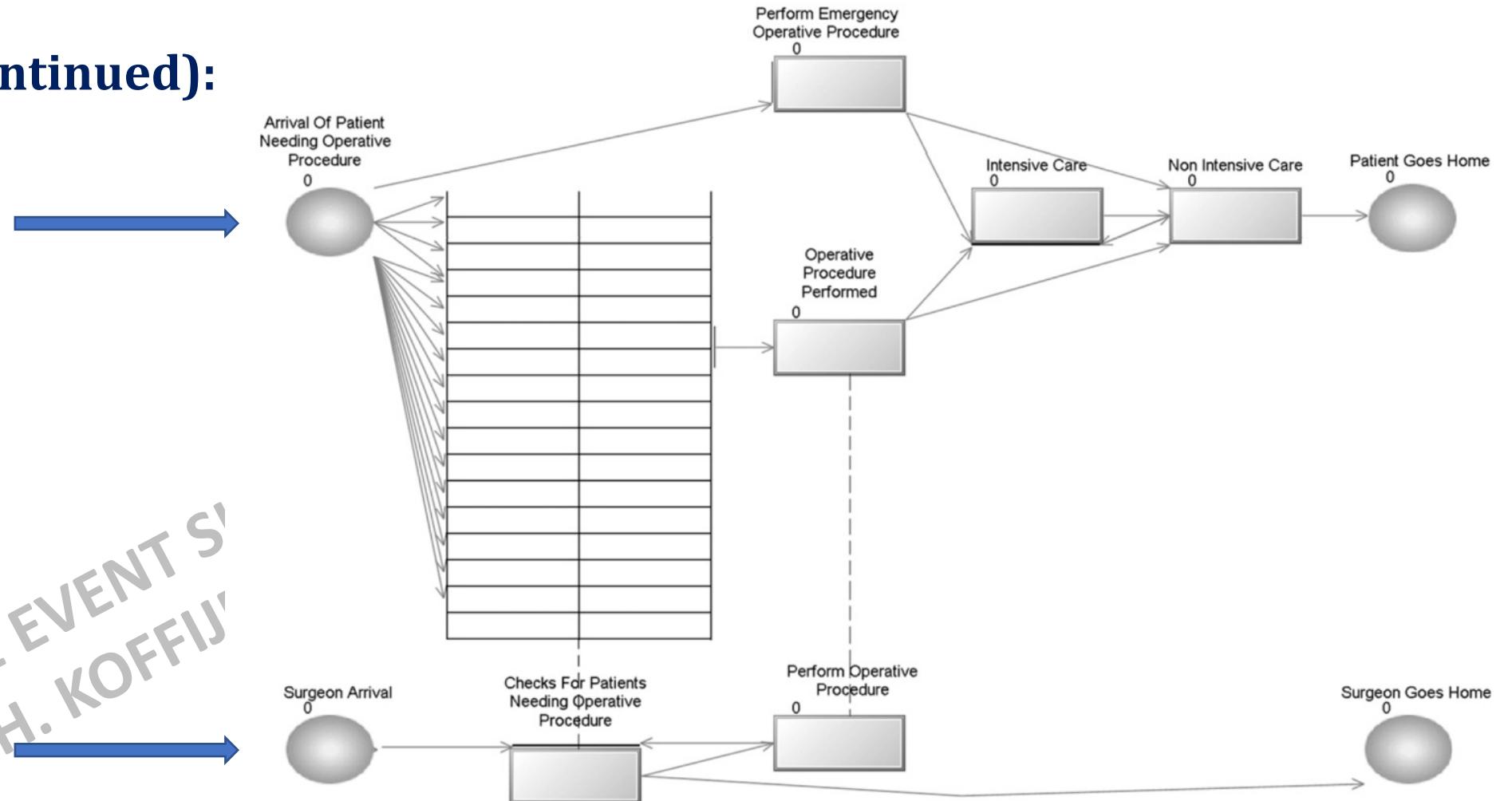


Figure. Visual representation of simulation model.

# Applications in healthcare

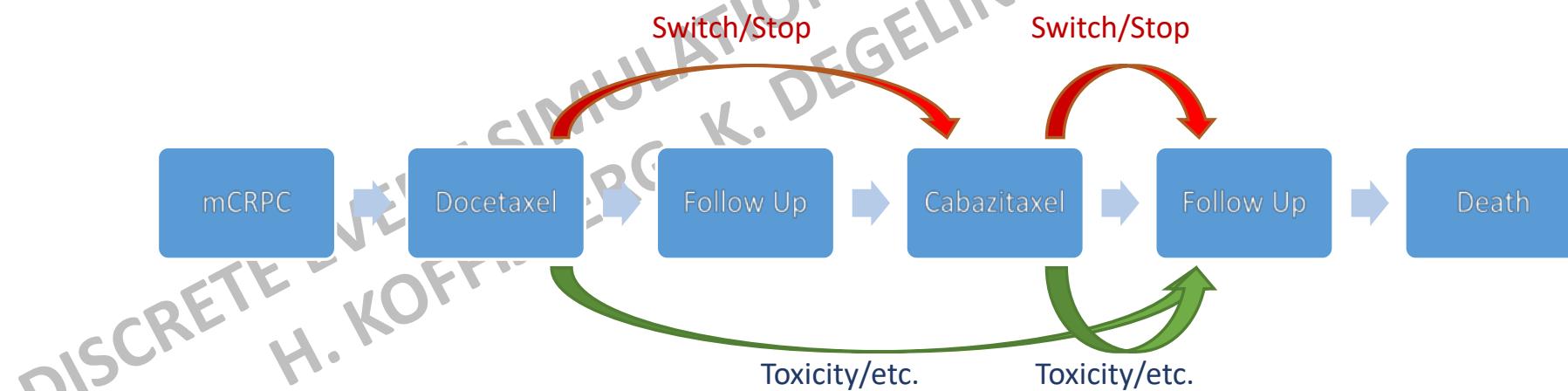
- Example (continued):
  - Performance measures:
    - Average number of days patients wait for cardiac interventions known to require an ICU stay
    - Number of cardiac operative procedures, known to require an ICU stay, that were actually performed
    - Average number of patients that should have been in the ICU, but were instead parked elsewhere
    - Number of blocks in which at least one intervention was cancelled because of insufficient functional ICU capacity
    - Utilization of ICU beds

# Applications in healthcare

- Example (continued):
  - Results and Conclusions
    - More functional ICU capacity leads to better performance:
    - Mean days until intervention
    - Average number of parked ICU patients
    - Mean number of surgical block cancellations
    - Mean number of cardiac interventions
  - However, performance curves “flatten out” with increasing capacity
  - Utilization of total ICU capacity is not an increasing function of ICU capacity
  - Hospital management needs to interpret results -> understanding of system

# Applications in healthcare

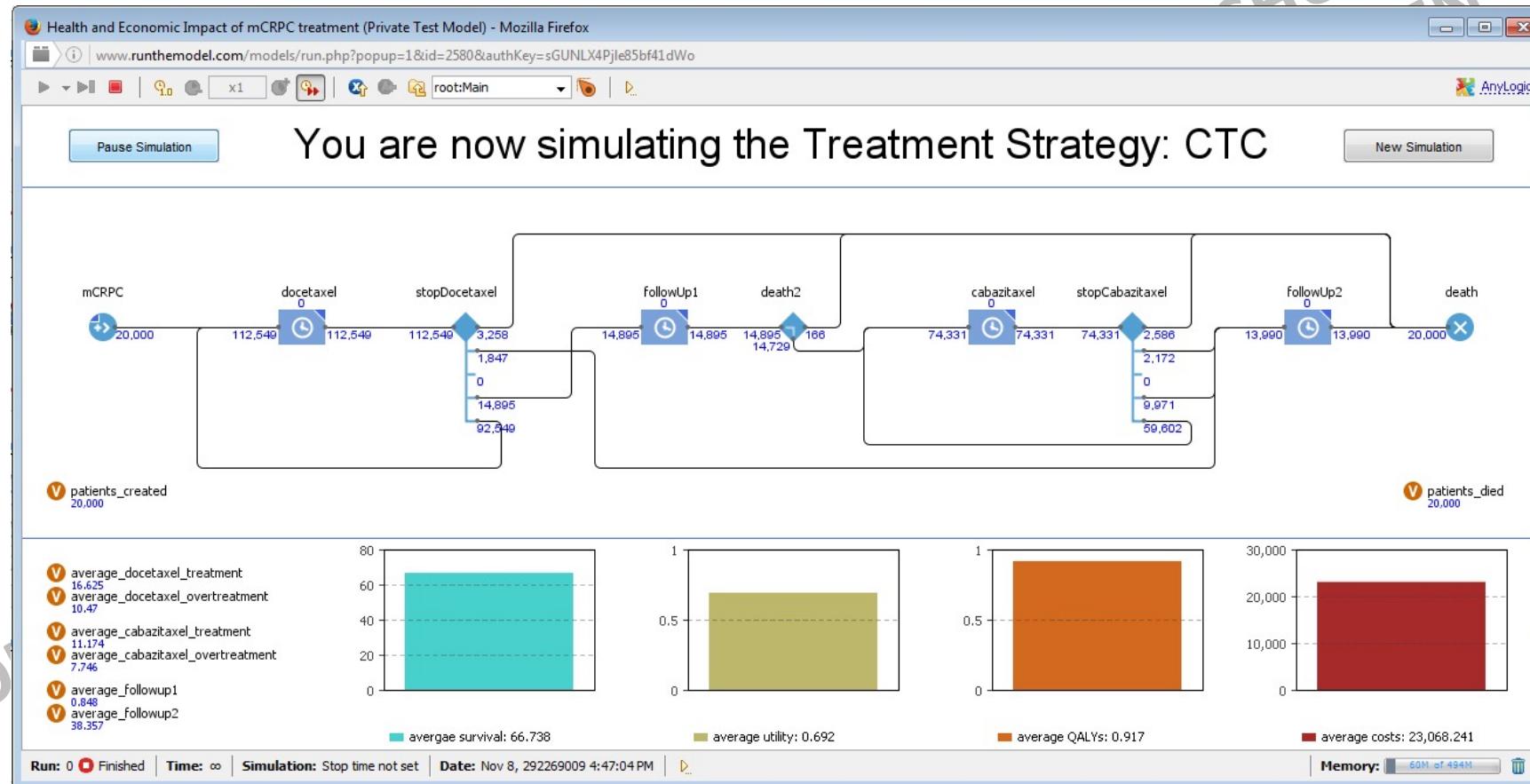
- Example:
  - Decision tool: circulating tumor cell enumeration in metastatic castration-resistant prostate cancer treatment (mCRPC)
    - Treatment monitoring using liquid biopsies (CTCs) as a response marker to guide treatment switches



Degeling K, et al. Comparison of Timed Automata with Discrete Event Simulation for Modeling of Biomarker-Based Treatment Decisions: An Illustration for Metastatic Castration-Resistant Prostate Cancer. Value in Health. 2017;20(10):1411-9.

# Applications in healthcare

- Example (continued):



# Lecture wrap-up

1. The basics of time-to-event-modeling
2. DES compared to other model types
3. An overview of potential applications in healthcare



DISCRETE EVENT SIMULATION – SMDA 2021 SHORT COURSE  
H. KOFFIJBERG, K. DEGELING, M. VAN DE VEN

# Discrete Event Simulation in R to Support Healthcare Decision Making

Modeling Competing Events in DES

**H. Koffijberg, PhD, MSc**

University of Twente  
Enschede, The Netherlands

**K. Degeling, PhD, MSc**

University of Melbourne  
Melbourne, Australia

**M. van de Ven, MSc**

OPEN Health  
Rotterdam, The Netherlands



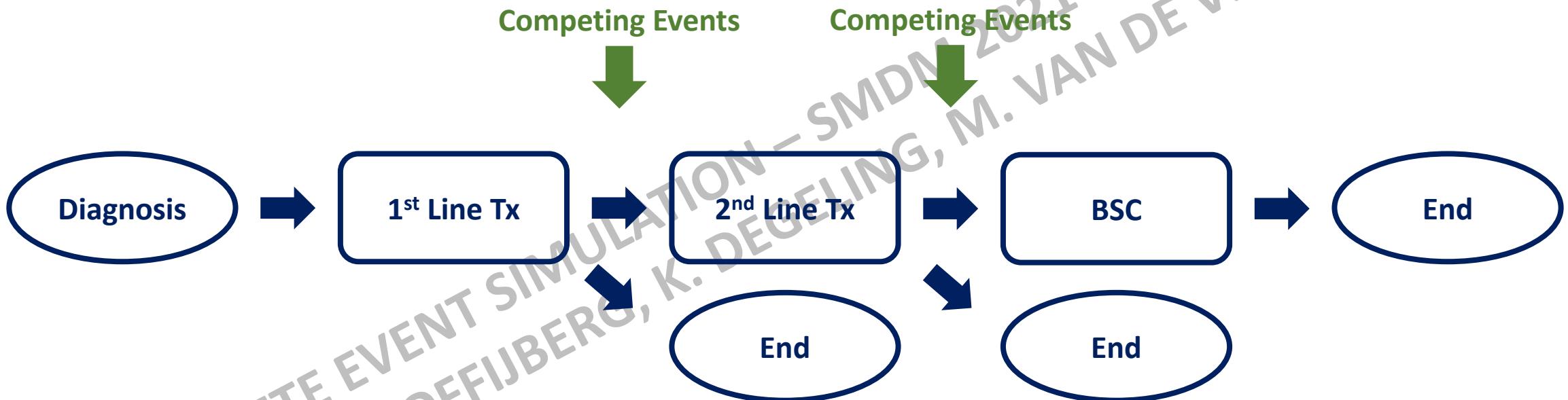
# Modeling Competing Events

1. What are Competing Events?
2. Competing Events in DES
3. Approach 1: First Event to Occur
4. Approach 2: Event First, Time Second
5. Approach 3: Time First, Event Second
6. Selecting Between Approaches

# Competing events

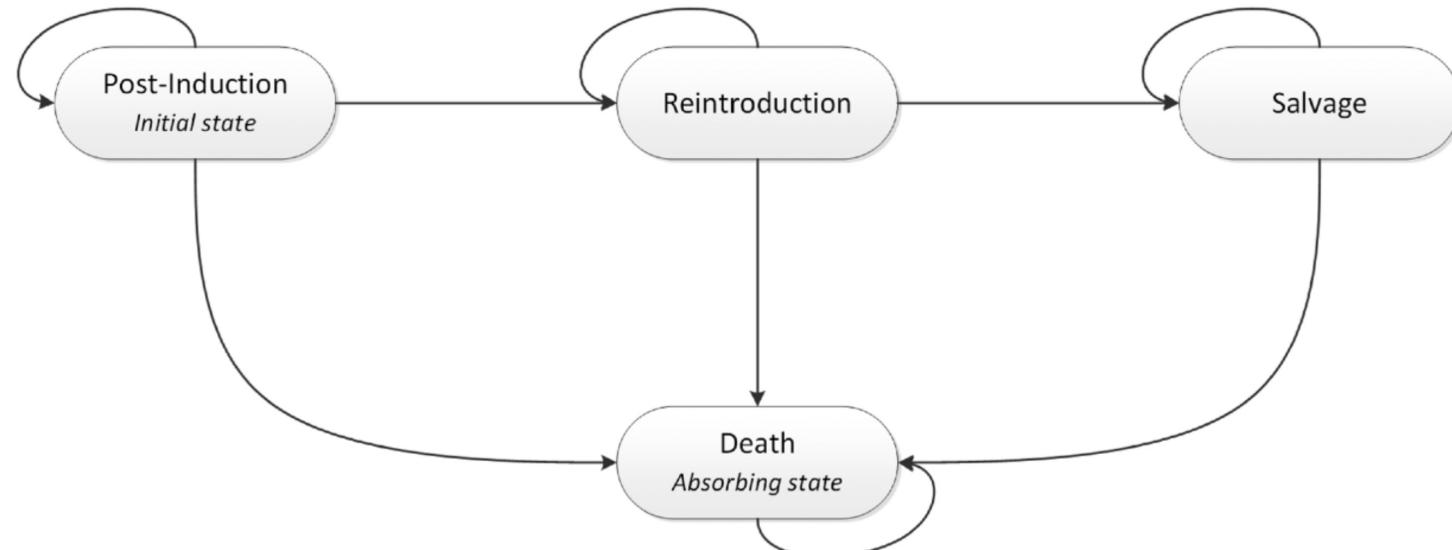
- Competing events either hinder the observation of the (other) event(s) of interest or modify the chance that these events occur
- Common in medical research:
  - Recurrence-free survival: *recurrence vs. death due to other causes*
  - Progression-free survival: *progression vs. death due to other causes*
  - Disease-specific survival: *death due to disease vs. due to other causes*
- Even more common in decision-analytic modeling...

# Relevance to decision-analytic modeling

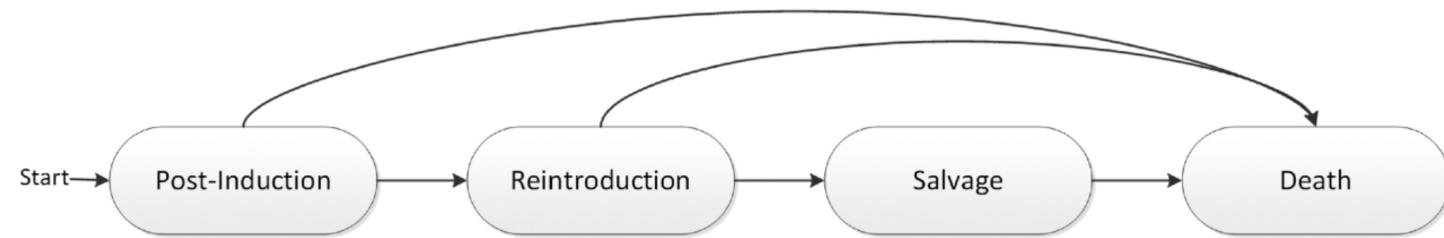


# State-transition model vs. DES

Discrete-time  
state-transition  
model:

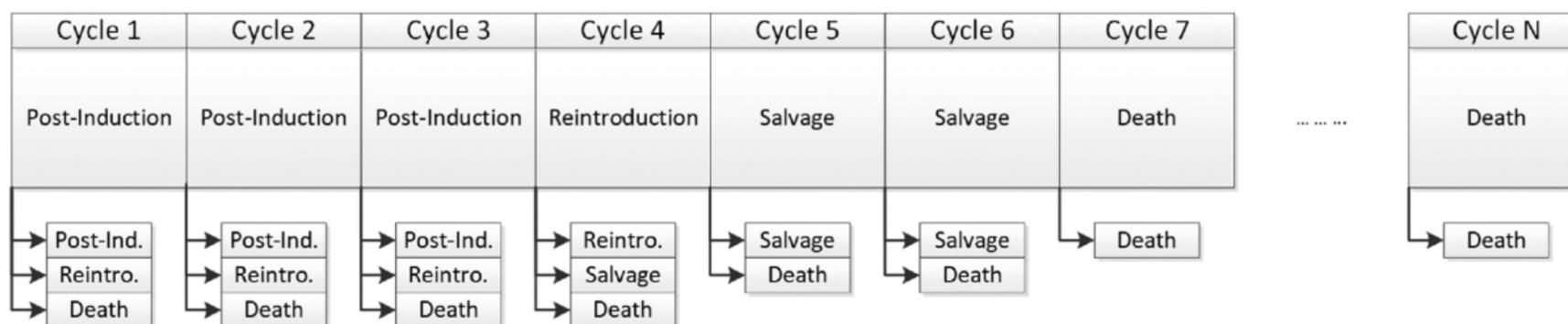
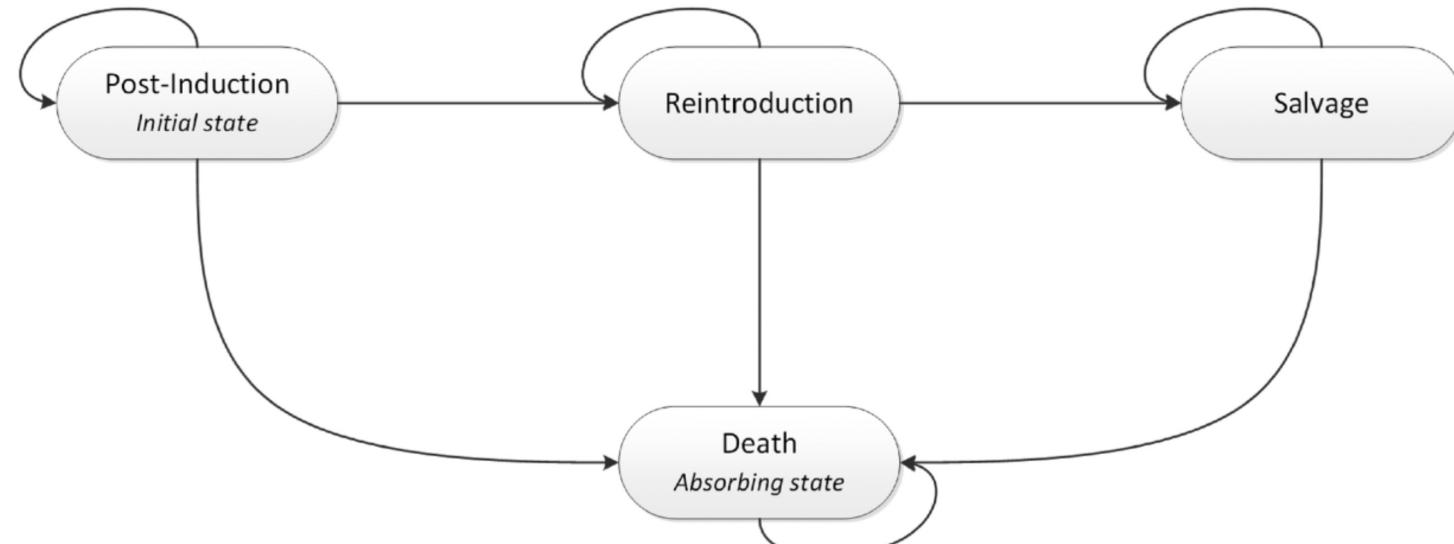


DES:



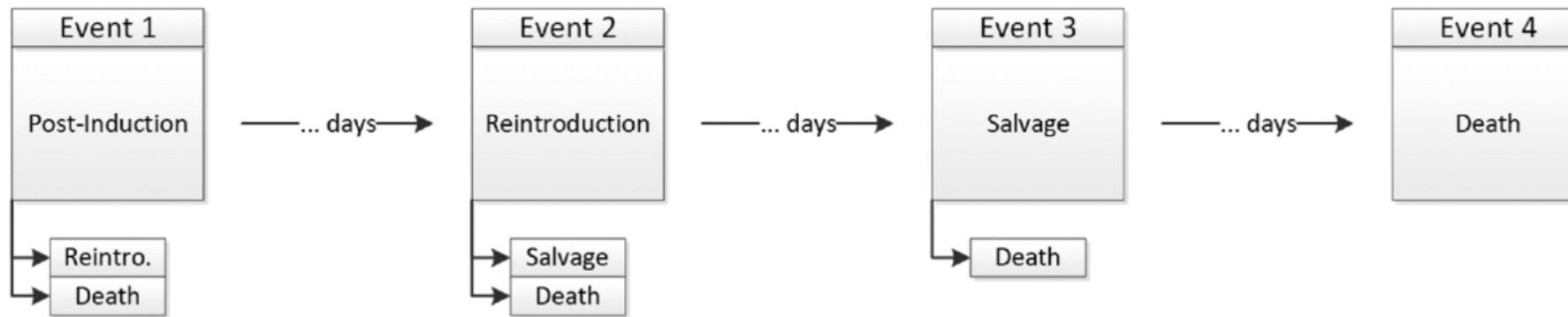
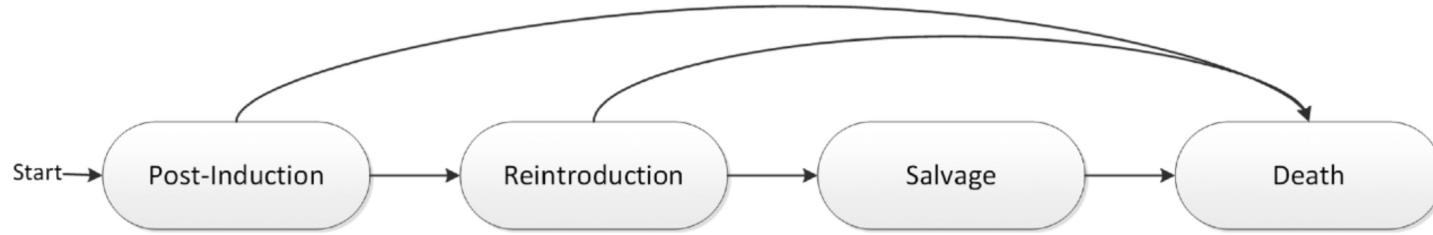
# State-transition model vs. DES

Discrete-time  
state-transition  
model:



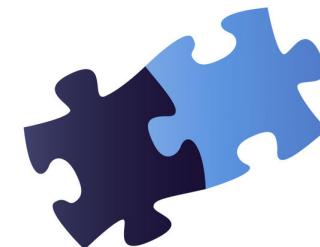
# State-transition model vs. DES

DES:



# Modeling approaches

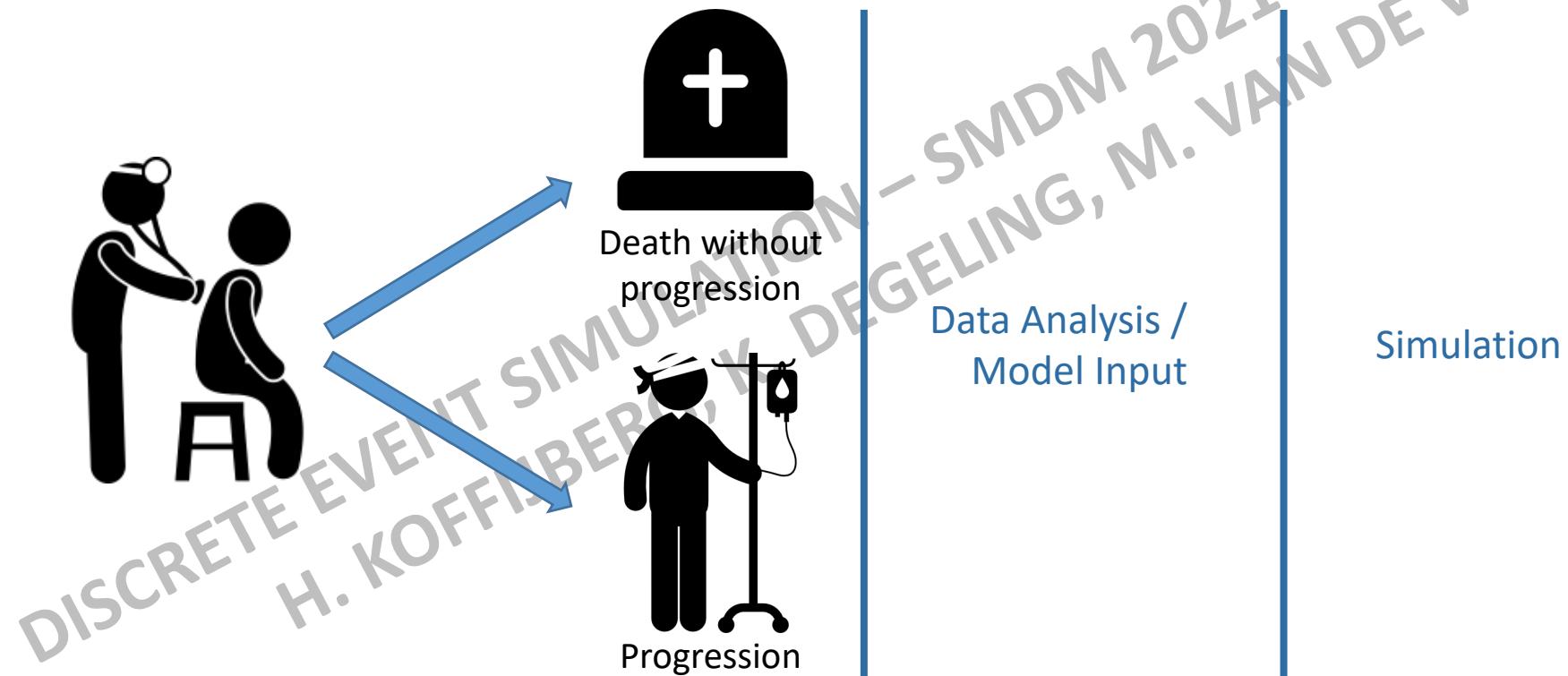
- Access to individual-level data:
  1. Sample times to each event and select the first to occur
  2. Sample the event first and the time-to-event second
  3. Sample the time-to-event first and the event second
  4. Based on discrete-time transition probabilities
    - Resembles a state-transition model
- Modeling approach = data analysis + simulation
- Aggregate data (e.g., from literature):
  - Already analyzed, so simulation should match the data analysis



Barton et al (2004): *The use of modelling to evaluate new drugs for patients with a chronic condition: the case of antibodies against tumour necrosis factor in rheumatoid arthritis.*

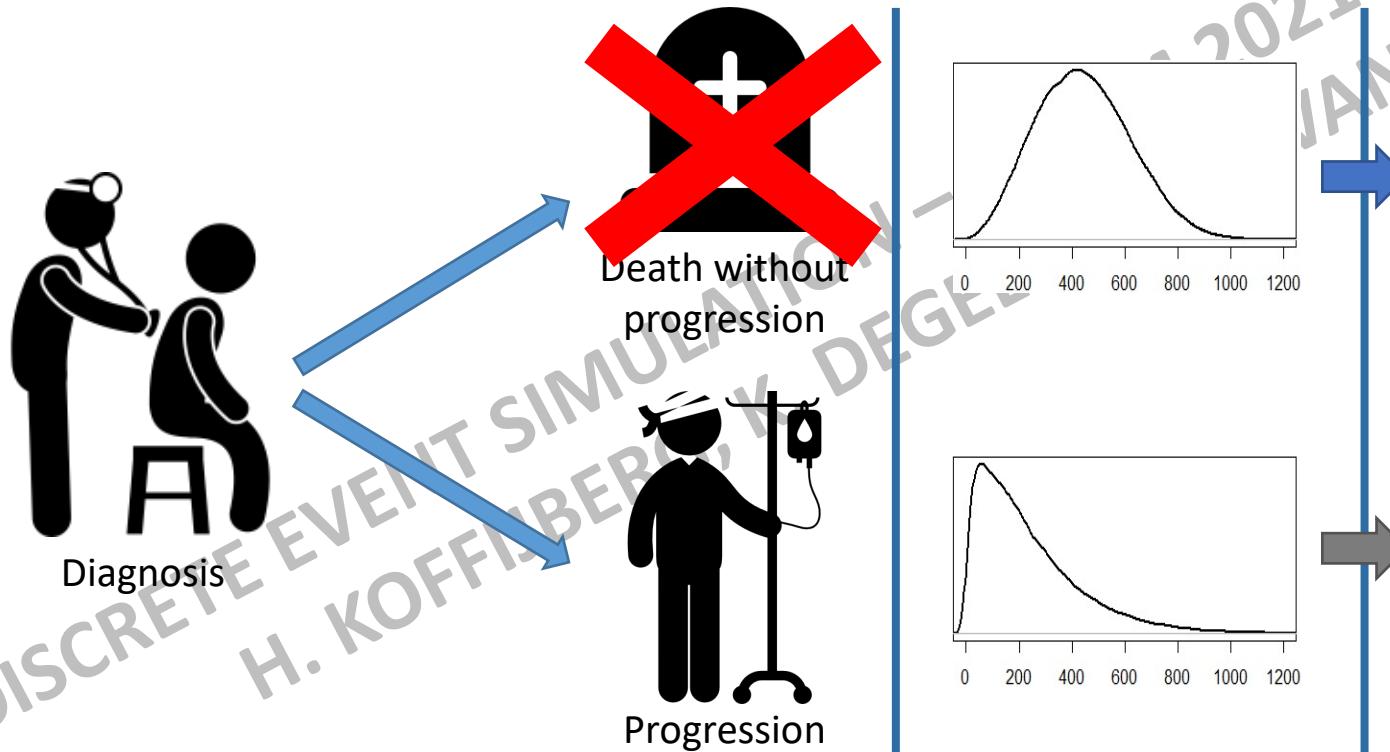
Degeling et al (2019): *Comparing strategies for modeling competing risks in discrete event simulations: a simulation study and illustration in colorectal cancer.*

# Illustration of modeling approaches



# Approach 1: first event to occur

## *Event-specific distributions (ESD)*



1) Sample time

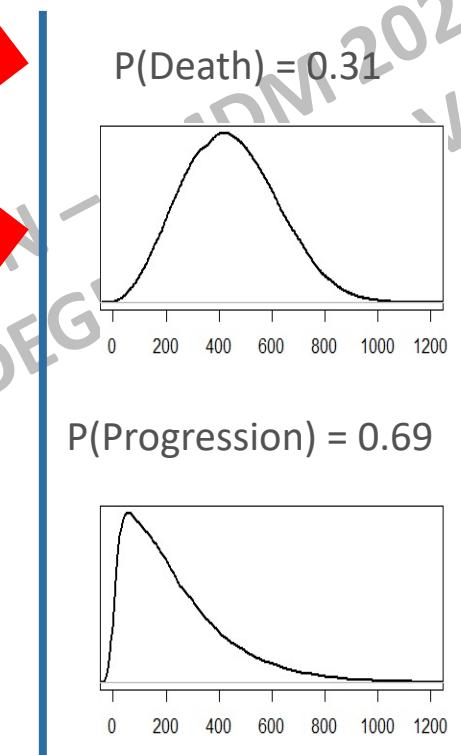
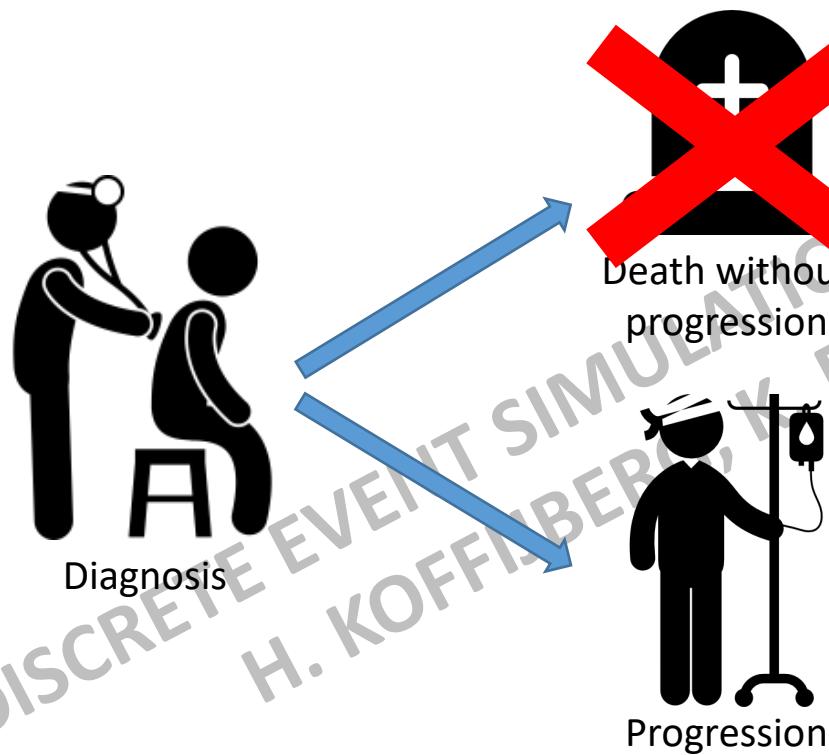
Random draw:  
250 days

2) Sample time

Random draw:  
93 days

# Approach 2: event first, time second

*Event-specific probabilities and distributions (ESPD)*



1) Sample event

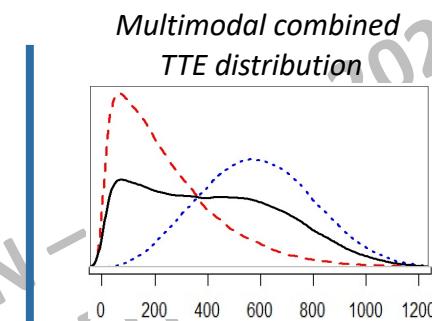
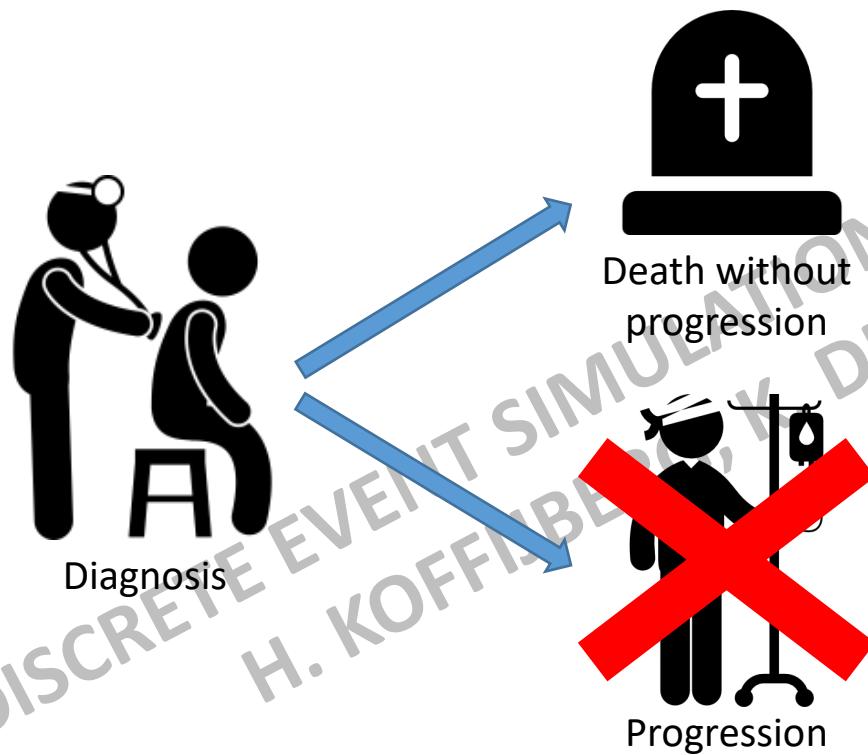
Random  $\sim U(0,1)$ :  
0.438

2) Sample time

Random draw:  
183 days

# Approach 3: time first, event second

*Unimodal / multimodal distribution and regression (UDR / MDR)*

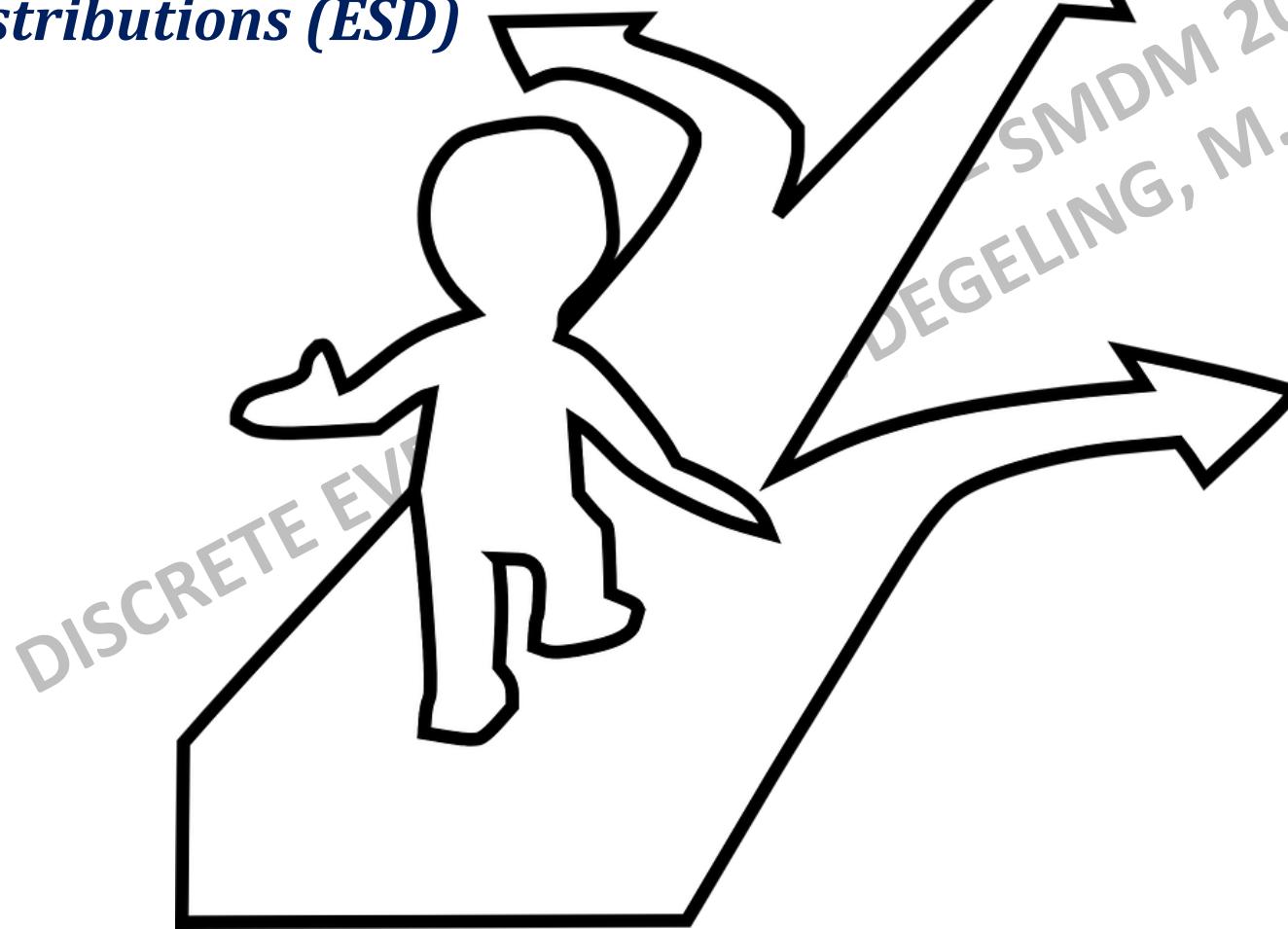


Regression model:  
 $Event \sim Time$

- 1) Sample time  
Random draw:  
 $274 \text{ days}$
- 2) Predict probabilities  
 $P(\text{Death}) = 0.81$   
 $P(\text{Progression}) = 0.19$
- 3) Sample event  
Random  $\sim U(0,1):$   
 $0.612$

# Selecting a modeling approach

*Event-specific  
distributions (ESD)*



*Event-specific probabilities  
and distributions (ESPD)*

*Unimodal / multimodal  
distribution and regression  
(UDR / MDR)*

# Aggregate vs. individual-level data

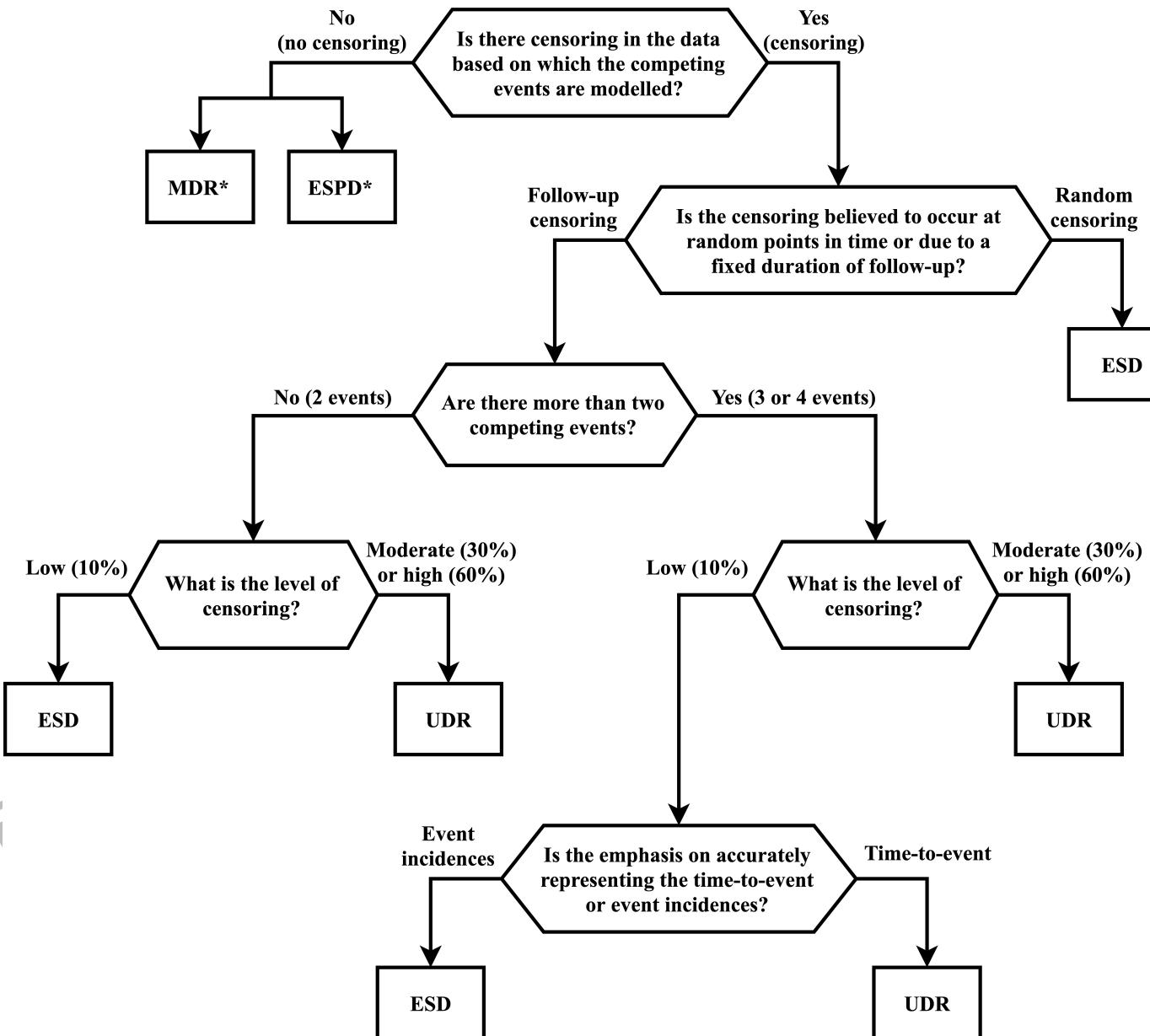
- When using aggregate data, the way that data has been analysed dictates what modelling approach is to be used
  - Also when combining with individual-level data (e.g., using a hazard ratio)
  - Generally, the ESD approach (i.e., first-event to occur)
- When using individual-level data, data characteristics may guide what modelling approach is suitable
  - Sample size / number of observed events
  - Number of competing events
  - Overlap of time-to-event distributions (i.e., multimodality)
  - Presence, type, and level of censoring

# Uncensored individual-patient data

- All three modelling approaches can be implemented
- Qualitative reasons may guide selection
  - Observational data
  - Heterogeneity
- ESPD and MDR preferred in terms of performance
  - With ESPD being the most straightforward to implement

# Censored individual-patient data

- ESD and UDR most straight forward
  - MDR depends on level of censoring
- Qualitative reasons may guide selection
- In terms of performance
  - Type of censoring
  - Number of competing events
  - Overlap of time-to-event distributions
  - Level of censoring
  - Importance of event incidence vs. time-to-event



# Awareness and validation are key

- Guidance is helpful, but every modelling challenge is different
- Chance of each event occurring
  - Event probabilities
- Time-to-event distribution(s)
  - Summary characteristics
  - Kaplan-Meier curves
  - Distribution tails

# Discrete Event Simulation in R to Support Healthcare Decision Making

Fitting distributions to patient-level and aggregate data

**H. Koffijberg, PhD, MSc**

University of Twente  
Enschede, The Netherlands

**K. Degeling, PhD, MSc**

University of Melbourne  
Melbourne, Australia

**M. van de Ven, MSc**

OPEN Health  
Rotterdam, The Netherlands

# Probability distributions in health economics

- Useful to reflect all kinds of variation / uncertainty
- Distinguish between distributions reflecting a value for a group of individuals vs reflecting variation in individuals' times, outcomes

# Probability distributions in health economics

- Some common parametric distributions for reflecting variation in
  - Probabilities (group level): Beta, Dirichlet
  - Relative risks (group level): Lognormal
  - Cost values (individual and group level): Gamma, Lognormal
  - Utility values (individual level): Beta, Gamma
  - Time-to-event (individual level): Weibull, Gompertz, Lognormal

Avoid discontinuous distributions

- Uniform
- Triangular

# Fitting probability distributions in R

Library(fitdistrplus): fitdist function

- Fitting on empirical data, using MLE, Moment Matching, GoF
- Visual comparison (densities, QQ-plot, PP-plot, etc.)
- Best fit based on statistical comparison: AIC and BIC
- Extensive examples in the package documentation

Library(rriskDistributions): fit.cont/fit.perc functions

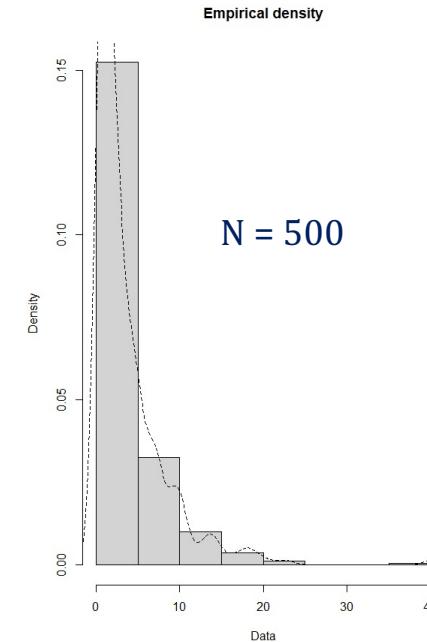
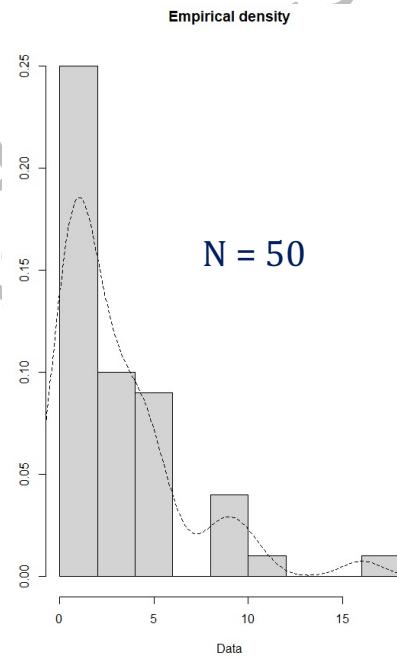
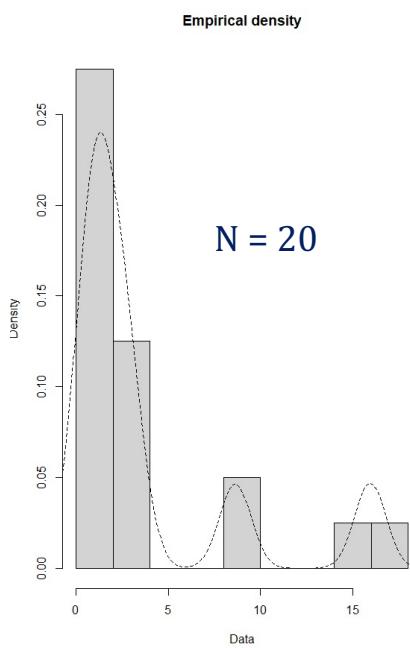
- Fitting on empirical data or known quantile values
- Fit up to 15 distributions at once
- Compare fit visually and numerically
- Select the best fit manually

# Defining probability distributions

- When IPD is available - *straightforward*
  - Use empirical distribution (IPD sufficiently large and smooth)
  - Fit a parametric distribution (limited IPD)
  - Reflect uncertainty in parameters defining this distribution
- When only aggregate data is available (common) - *bit harder*
  - Define parametric distribution based on reported **mean, SD, SE, sample size**
- Without any evidence, using expert opinion - *more work*
  - Ask an expert panel to estimate variation on individual level, pool the resulting estimates (after weighting if relevant)

# Defining the empirical distribution- patient-level data

- Using the distribution in the empirical data
- The sample size needs to be sufficiently large to result in a smooth distribution



# Defining probability distributions – aggregate data

- Recalculate reported *mean, SD, SE, sample size* to parameters of the target distribution
  - One technique: **Method of Moments**
  - A moment is a specific quantitative measure of the shape of a function.
  - The zeroth moment is the total probability (i.e. one)
  - The first moment is the expected value,
  - The second central moment is the variance
  - The third standardized moment is the skewness
  - Target distributions with 2 parameters can be defined (parametrized) when you know the value of the first and second moment

# Defining probability distributions – aggregate data

$$\Theta \sim \text{Beta}(\alpha, \beta)$$

Mean:  $E[\Theta] = \alpha/(\alpha+\beta)$  suppose the known mean is  $\bar{u}$

Variance:  $\text{Var}[\Theta] = \alpha\beta/((\alpha+\beta)^2 * (\alpha+\beta+1))$  suppose the known variance is  $s^2$

$$\bar{u} = \alpha / (\alpha+\beta)$$

>>> insert mathematics here >>>

$$(\alpha+\beta) = \bar{u} (1 - \bar{u})/s^2 - 1$$

$$s^2 = \alpha\beta / (\alpha+\beta)^2(\alpha+\beta+1)]$$

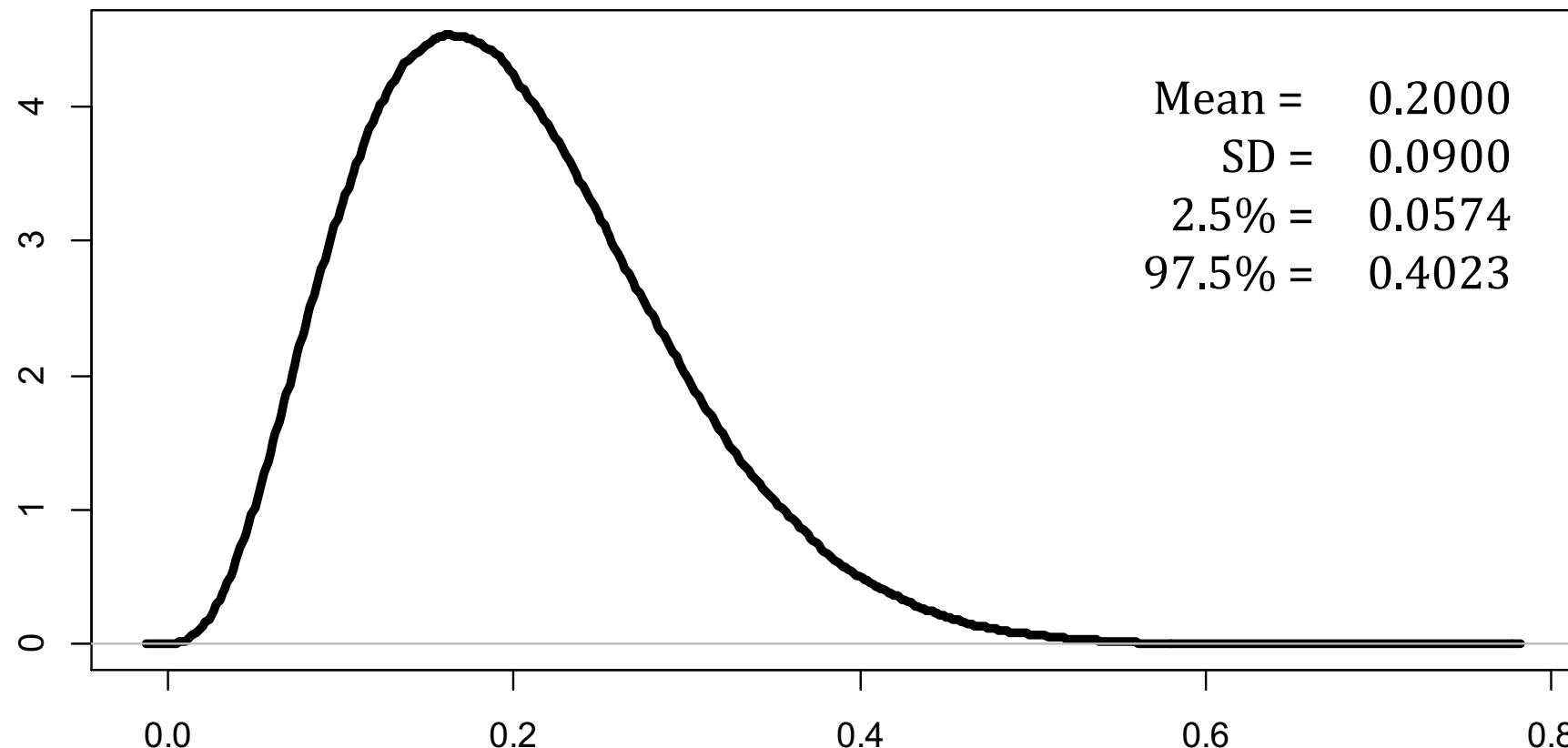
$$\alpha = \mu(\alpha+\beta)$$

*Example for observed data:  $\bar{u} = 0.20$ ,  $sd = s = 0.09$*

$$(\alpha+\beta) = 0.20 * (1 - 0.20)/0.09^2 - 1 \approx 18.75$$

$$\alpha = 0.20 * 18.75 = 3.75, \text{ and then of course } \beta = 18.75 - 3.75 = 15.00$$

# Defining probability distributions – aggregate data



*Example for observed data:  $\bar{u} = 0.20$ ,  $sd = s = 0.09$*

$$(\alpha + \beta) = 0.20 * (1 - 0.20) / 0.09^2 - 1 \approx 18.75$$

$$\alpha = 0.20 * 18.75 = \mathbf{3.75}, \text{ and then of course } \beta = 18.75 - 3.75 = \mathbf{15.00}$$

# Defining probability distributions – aggregate data

$\Theta \sim \text{Gamma}(\alpha, \beta)$

Mean:  $E[\Theta] = \alpha/\beta$

Variance:  $\text{Var}[\Theta] = \alpha/\beta^2$

suppose the known mean is  $\bar{u}$

suppose the known variance is  $s^2$

$$\bar{u} = \alpha/\beta$$

$$s^2 = \alpha/\beta^2$$

$$\alpha = \bar{u}^2 / s^2$$

$$\beta = \bar{u} / s^2$$

>>>insert mathematics here >>>

**NOTE** for first order uncertainty  $s^2 = SD^2$  of sample  
for second order uncertainty  $s^2 = SE^2$  of mean

# Approximating IPD from aggregate data

- Applicable to time-to-event data
- Useful for HE modeling and meta-analyses
- Process:
  - Use digitizing software to retrieve the survival probabilities, numbers at risk, and total number of events from KM plots
  - Then use one of the published algorithms to approximate IPD
    - Most well-known: Guyot et al. (2012)
- Review of Wan et al. (2015) compared differences between methods
  - Conclusion: Best choice depends on target distributions

Guyot, P., Ades, A. E., Ouwens, M. J., & Welton, N. J. (2012). Enhanced secondary analysis of survival data: reconstructing the data from published Kaplan-Meier survival curves. *BMC medical research methodology*, 12(1), 1-13.

Wan, X., Peng, L., & Li, Y. (2015). A review and comparison of methods for recreating individual patient data from published Kaplan-Meier survival curves for economic evaluations: a simulation study. *PLoS One*, 10(3), e0121353.

# Questions

DISCRETE EVENT SIMULATION  
H. KOFFIJBERG, K. DEGELING, M. VAN DE VEN  
SMDM 2021 SHORT COURSE

# Discrete Event Simulation in R to Support Healthcare Decision Making

Uncertainty

**H. Koffijberg, PhD, MSc**

University of Twente  
Enschede, The Netherlands

**K. Degeling, PhD, MSc**

University of Melbourne  
Melbourne, Australia

**M. van de Ven, MSc**

OPEN Health  
Rotterdam, The Netherlands

# Topics

- Introduction & stochastic uncertainty
- Parameter uncertainty & probabilistic analysis
- Summary

DISCRETE EVENT SIMULATION – SUMMER 2021 SHORT COURSE  
H. KOFFIJBERG, K. DEGELING, M. VAN DE VEN

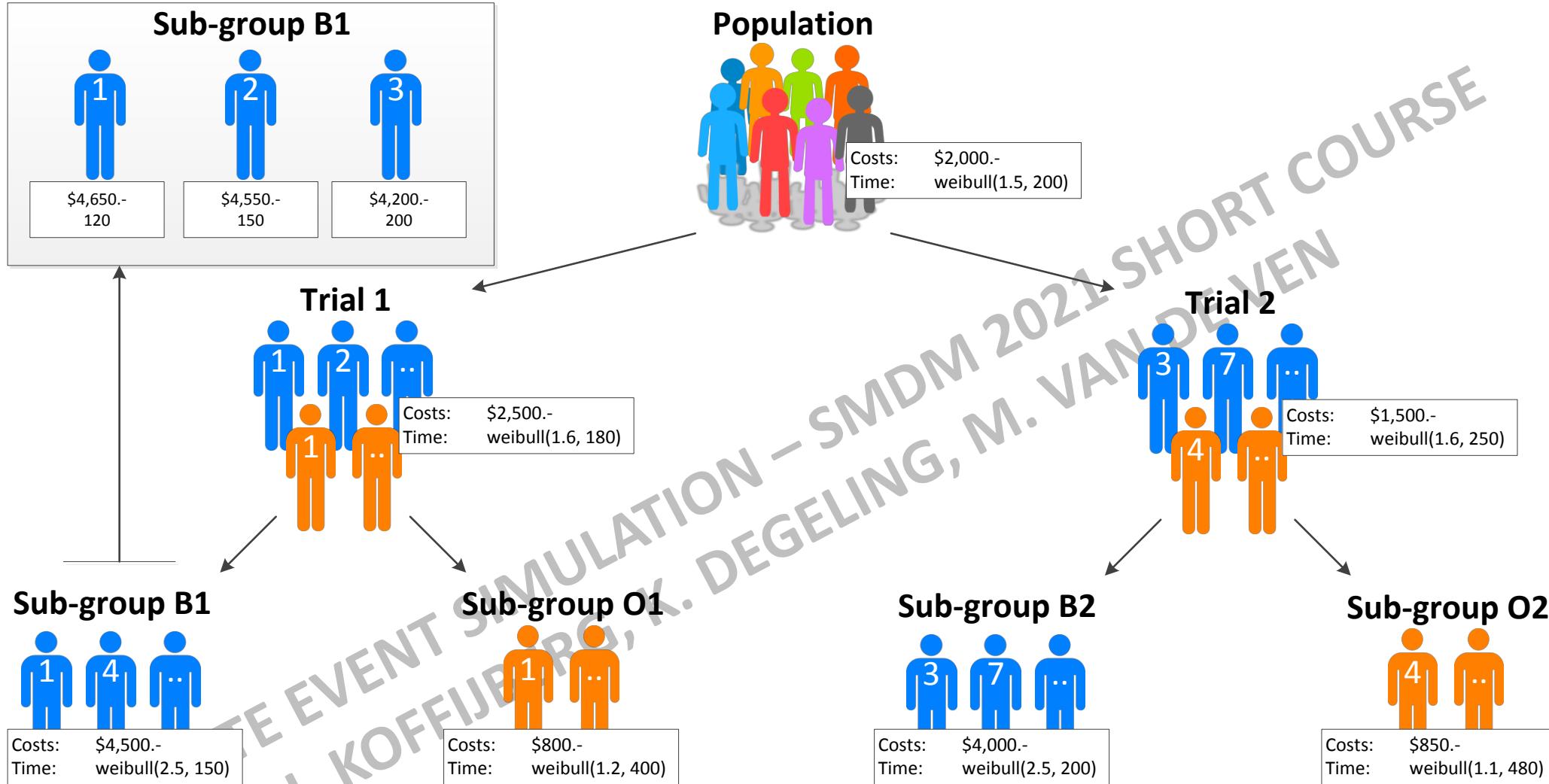
# Introduction & Stochastic uncertainty

- Introducing the different types of uncertainty
- Understanding the role of stochastic uncertainty in DES
- Providing guidance for determining the number of entities to be simulated

# Types of uncertainty

Preferred term	Concept	Other terms used	Analogous concept in regression
----------------	---------	------------------	---------------------------------

Briggs AH, et al. Model Parameter Estimation and Uncertainty: A Report of the ISPOR-SMDM Modeling Good Research Practices Task Force-6. Value Health. 2012;15:835-42



Differences in outcomes between **trials** due to:

Differences in outcomes between **subgroups between trials (B1 & B2)** due to:

Differences in outcomes between **subgroups in a trial (B1 & O1)** due to:

Differences in outcomes between **patients in a subgroup** due to:

# Stochastic uncertainty

- Reflection of stochastic uncertainty is a key aspect of DES:
  - Through time-to-event distributions
  - Through event probabilities
  - Through distributions for outcomes
- If distributions are skewed (i.e. non-symmetrical), stochastic uncertainty will result in different mean outcome estimates
- Stochastic uncertainty needs to be “simulated out” by simulating sufficient entities to obtain stable mean outcome estimates
  - Applies to all patient-level modeling paradigms

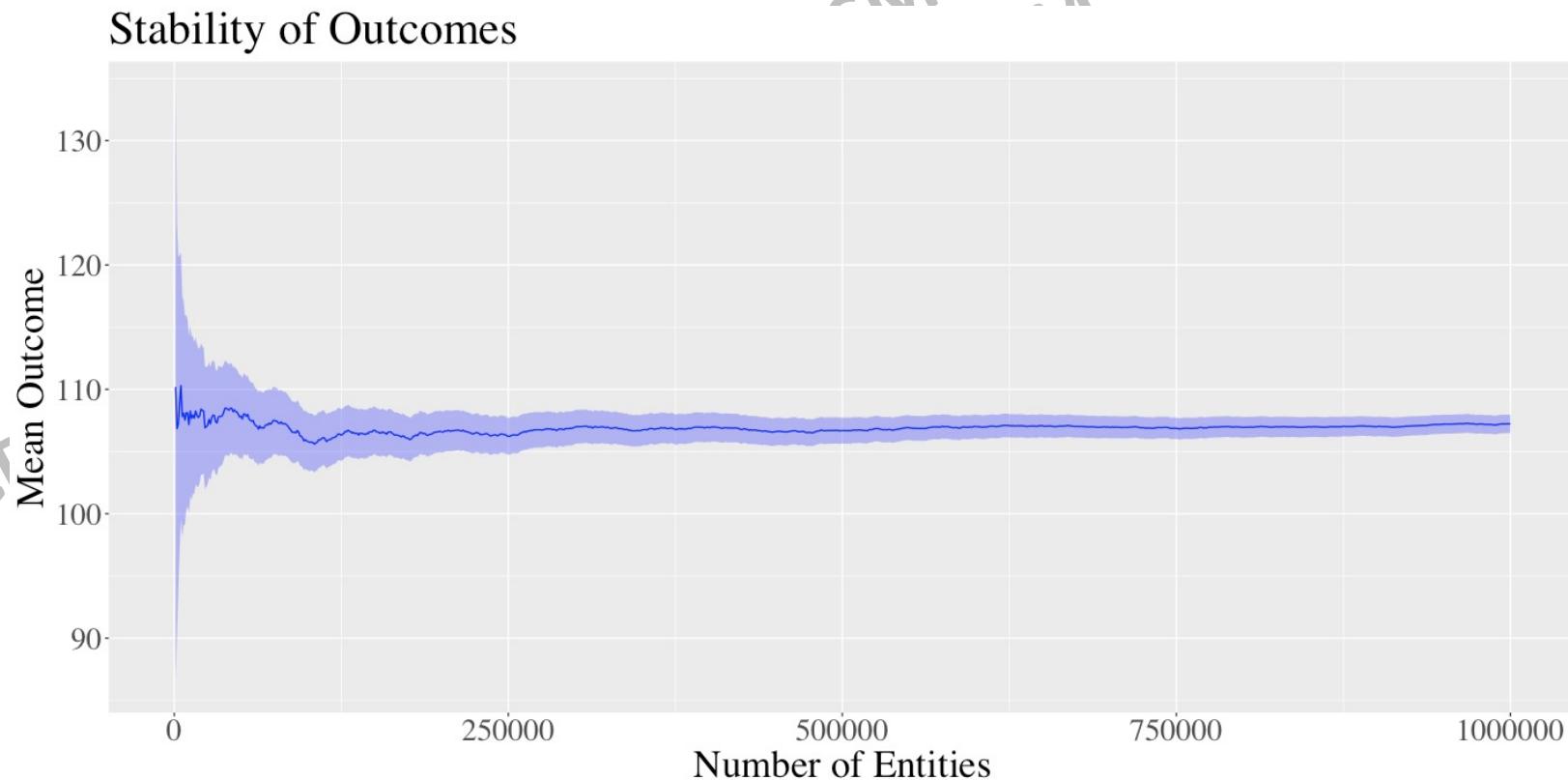
# Stochastic uncertainty

- How to determine the number of entities that need to be simulated to obtain stable outcome estimates?
  1. Visual inspection
  2. Mathematical estimation

	$n = 1,000$	$n = 10,000$	$n = 100,000$	$n = 1,000,000$
<b>Run 1</b>	108.16	104.69	106.02	107.23
<b>Run 2</b>	110.08	100.70	107.47	106.11
<b>Run 3</b>	99.02	98.20	106.41	106.49
<b>Run 4</b>	109.78	102.43	106.71	106.67

# Stochastic uncertainty

- Number of entities to be simulated: *visual inspection*
  1. Perform one simulation run with a relatively high number of entities
  2. Calculate the mean outcome for an increasing proportion of the entities
  3. Plot the mean outcome vs. number of entities



# Stochastic uncertainty

- Number of entities to be simulated: *mathematical estimation*
  - Run the model with a certain number of entities ( $n$ )
  - Calculate the *mean* and *standard error (SE)* for the outcomes
  - Determine the number of entities to be simulated ( $n^*$ ) to reach a pre-defined value for the *standard error (SE\*)*

	$n = 1,000$	$n = 10,000$
<b>Mean</b>	108.16	104.69
<b>SD</b>	383.88	373.30
<b>SE</b>	12.14	3.73
<b>95% LB</b>	84.37	97.37
<b>95% UB</b>	131.95	112.01

$$n^* \approx \left( \frac{SE}{SE^*} \right)^2 \times n$$

Goal:  $SE^* = 0.01 \times 105 = 1.05$

$$n^* \approx \left( \frac{3.73}{1.05} \right)^2 \times 10000 \approx 126,399$$

# Stochastic uncertainty

- Number of entities to be simulated: *mathematical estimation*
  - Run the model with a certain number of entities ( $n$ )
  - Calculate the *mean* and *standard error (SE)* for the outcomes
  - Determine the number of entities to be simulated ( $n^*$ ) to reach a pre-defined value for the *standard error (SE\*)*

	$n = 10,000$	$n = 130,000$
Mean	104.69	106.53
SD	373.30	377.97
SE	3.73	1.05
95% LB	97.37	104.47
95% UB	112.01	108.58

$$n^* \approx \left( \frac{SE}{SE^*} \right)^2 \times n$$

Goal:  $SE^* = 0.01 \times 105 = 1.05$

$$n^* \approx \left( \frac{3.73}{1.05} \right)^2 \times 10000 \approx 126,399$$

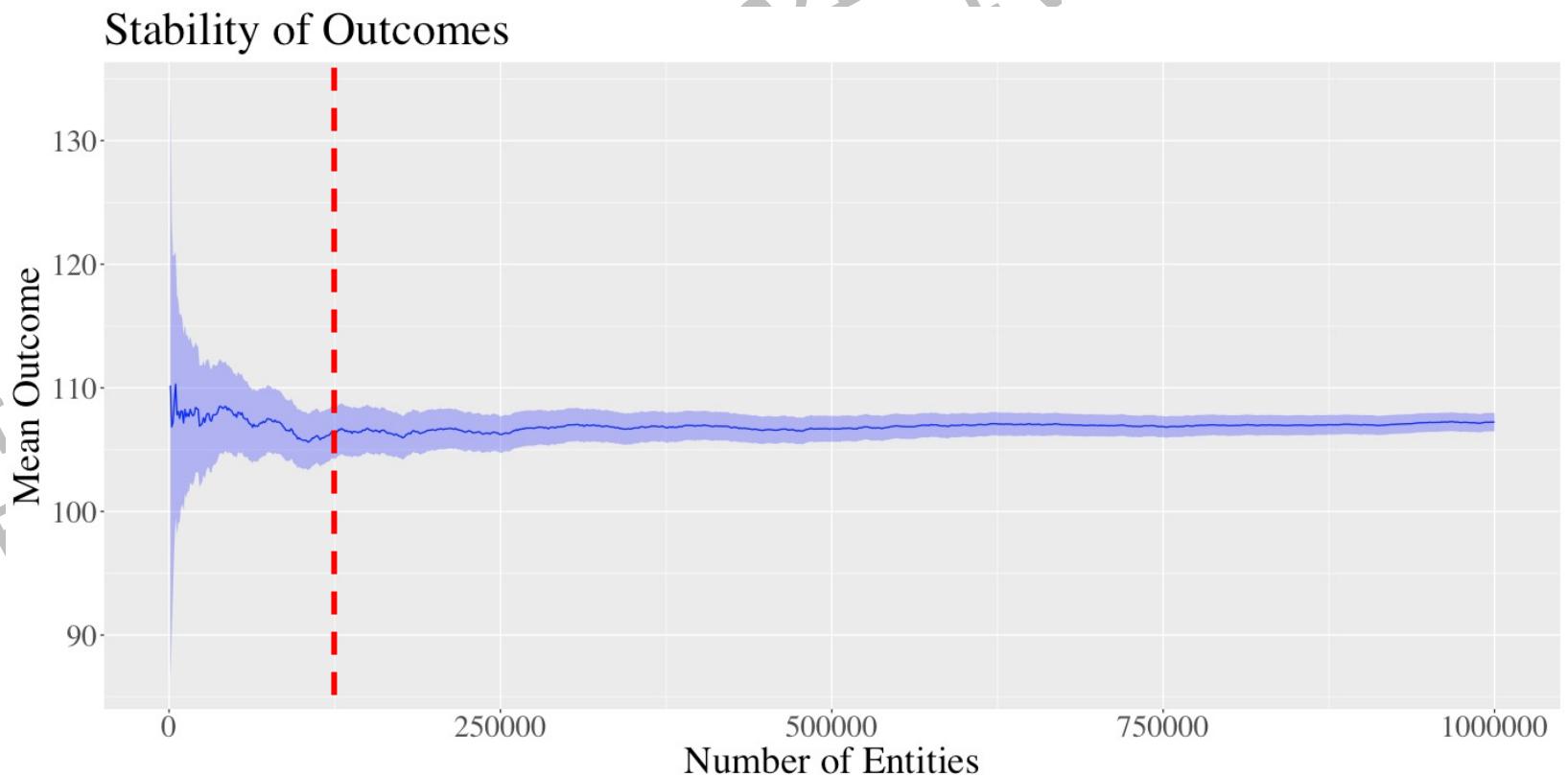
# Stochastic uncertainty

- Number of entities to be simulated: *comparison*

---

<b>n = 130,000</b>	
<b>Mean</b>	106.53
<b>SD</b>	377.97
<b>SE</b>	1.05
<b>95% LB</b>	104.47
<b>95% UB</b>	108.58

---



# Stochastic uncertainty

- Further aspects to consider:
  - Be careful calculating SD for differences and ratios
    - You cannot simply take the SD of patient-level differences/ratios!
    - Bootstrapping your simulated data is an option to get the SE
  - Differences and especially ratios tend to be very unstable
    - Differences: incremental outcomes
    - Ratios: incremental cost-effectiveness ratio (ICER), especially when the incremental effectiveness is small

# **Parametric uncertainty & probabilistic analysis**

- Understanding the role of parameter uncertainty in DES
- Providing guidance for reflecting parameter uncertainty in DES

# Parameter uncertainty

- Every model based analysis should reflect and synthesize the best available evidence at the moment of execution
- Most evidence stems from observations and measurements
  - Chance of progression in NSCLC patients during 1<sup>st</sup> line chemotherapy
  - Quality of life in patients experiencing side effects from chemotherapy
  - Cost of hospitalization due to side effects
- Most common exception
  - Fixed prices of technologies (drugs, E-Health apps, sensors)

**Almost all model parameters are based on *estimates* of an unknown, true underlying value**

# Parameter uncertainty

- Estimates always have uncertainty associated with them
- Reflecting this in health economic evaluations is done through

probabilistic sensitivity analysis

- Reflect uncertainty in combined evidence (second-order uncertainty)
- Defining and using probability distributions for all model parameters
- Type of probability distribution may vary per parameter, depending on parameter characteristics

# Parameter uncertainty

- Uncertainty in input (parameter values) should lead to uncertainty in outcomes, as a result of propagation through the model
- Uncertainty in outcomes matters because basing decisions about an intervention on expected outcomes ignores the question of whether current evidence is a *sufficient basis* for guiding decisions
  - Without sufficient/high quality evidence, adoption or reimbursement decisions about technologies will be uncertain: There will be a chance that the wrong decision is made resulting in a loss of health and resources

**ConVOI**

<https://www.convoi-group.org/>

**ISPOR VOI Taskforce reports 1 & 2**

- Fenwick et al. Value of Information Analysis for Research Decisions - An Introduction
- Rothery et al. Value of Information Analytical Methods

# Probabilistic analysis

© IJRSE

- 1) Define parametric/empirical probability distributions for all parameters
- 2) Replace mean estimates of cost, effect, utility parameters etc with distributions
  - DES = patient-level, therefore, “mean estimates” (also) relates to *parameters defining distributions* from which individual level values are drawn
- 3) Propagate uncertainty through the model
  - Draw a random value for each parameter from the corresponding distribution
  - Evaluate the model (repeat 5,000x)
- 4) Report the results
  - Distribution of outcomes per strategy/intervention
  - Chance that a strategy is acceptable (ICER)
  - Chance that a strategy is optimal

# Parameter uncertainty – DES specific approach

Degeling et al. BMC Medical Research Methodology (2017) 17:170  
DOI 10.1186/s12874-017-0437-y

BMC Medical Research  
Methodology

RESEARCH ARTICLE

Open Access



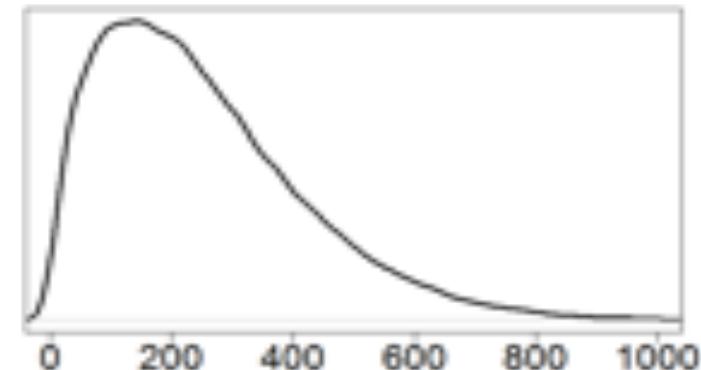
CrossMark

## Accounting for parameter uncertainty in the definition of parametric distributions used to describe individual patient variation in health economic models

D Koen Degeling<sup>1</sup> , Maarten J. IJzerman<sup>1</sup>, Miriam Koopman<sup>2</sup> and Hendrik Koffijberg<sup>1\*</sup>

# Parameter uncertainty – DES specific approach

Type of Uncertainty	Time-to-Event <i>(Stochastic uncertainty not reflected)</i>	Time-to-Event <i>(Stochastic uncertainty reflected)</i>
Stochastic uncertainty <i>(Patient-level variation)</i> <i>(First-order uncertainty)</i>	<b>Box A</b> <p>Patient-level variation is not described. Consequently, an estimated mean time-to-progression is assigned to all patients:</p> <ul style="list-style-type: none"><li>- Mean est. = 250 days</li></ul>	<b>Box C</b> <p>Patient-level variation in time-to-progression is described by a Weibull distribution:</p> <ul style="list-style-type: none"><li>- Shape parameter est. = 1.5</li><li>- Scale parameter est. = 277</li><li>- <math>\mu \approx 250</math> days</li></ul>



# Parameter uncertainty – DES specific approach

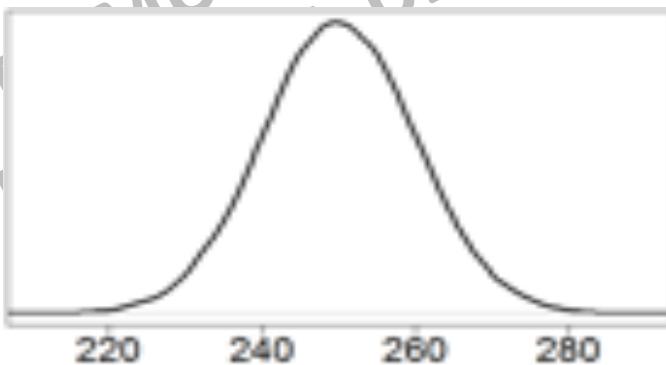
Parameter uncertainty  
*(Second-order uncertainty)*

## Box B

Parameter uncertainty in the estimated mean time-to-progression is described by a Normal distribution:

- Mean parameter est. = 250
- SEE parameter est. = 10

and included in the PSA.



## Box D

Parameter uncertainty in the estimated Shape and Scale parameter of the Weibull distribution needs to be described and included in the PSA. Patient-level variation in time-to-progression is reflected by drawing from a different distribution in each PSA sample:

PSA	Shape	Scale	Mean
1	1.50	242	219
2	1.62	253	227
3	1.67	292	260
n	...	...	...

# Parameter uncertainty – DES specific approach

Evaluation of 2 approaches in a simulation and case study

## 1) Using non-parametric bootstrapping

- Bootstrap the IPD to obtain a new estimate for the parameters  $(\alpha, \beta)$  of the target distribution
- Repeat the bootstrap N times, to support running the PA with N samples, and use different estimates for  $(\alpha, \beta)$  in every PA sample

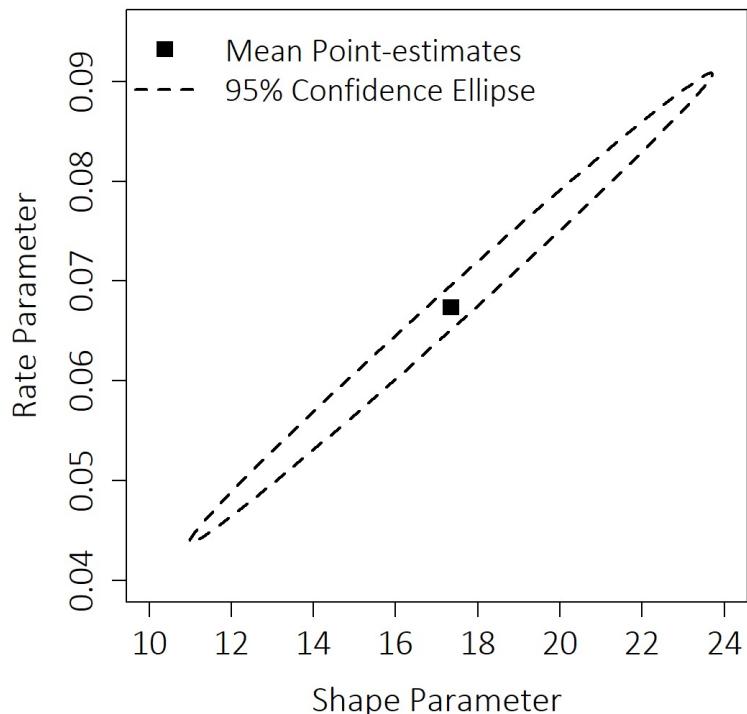
## 2) Using multivariate Normal distributions

- Estimate the parameters  $(\alpha, \beta)$  of the target distribution from the IPD
- Use the variance-covariance matrix of the estimates  $(\alpha, \beta)$  to define a multivariate normal distribution
- Draw N feasible sets of values for  $(\alpha, \beta)$  from the multivariate normal distribution, to support running the PA with N samples, and use different estimates for  $(\alpha, \beta)$  in every PA sample

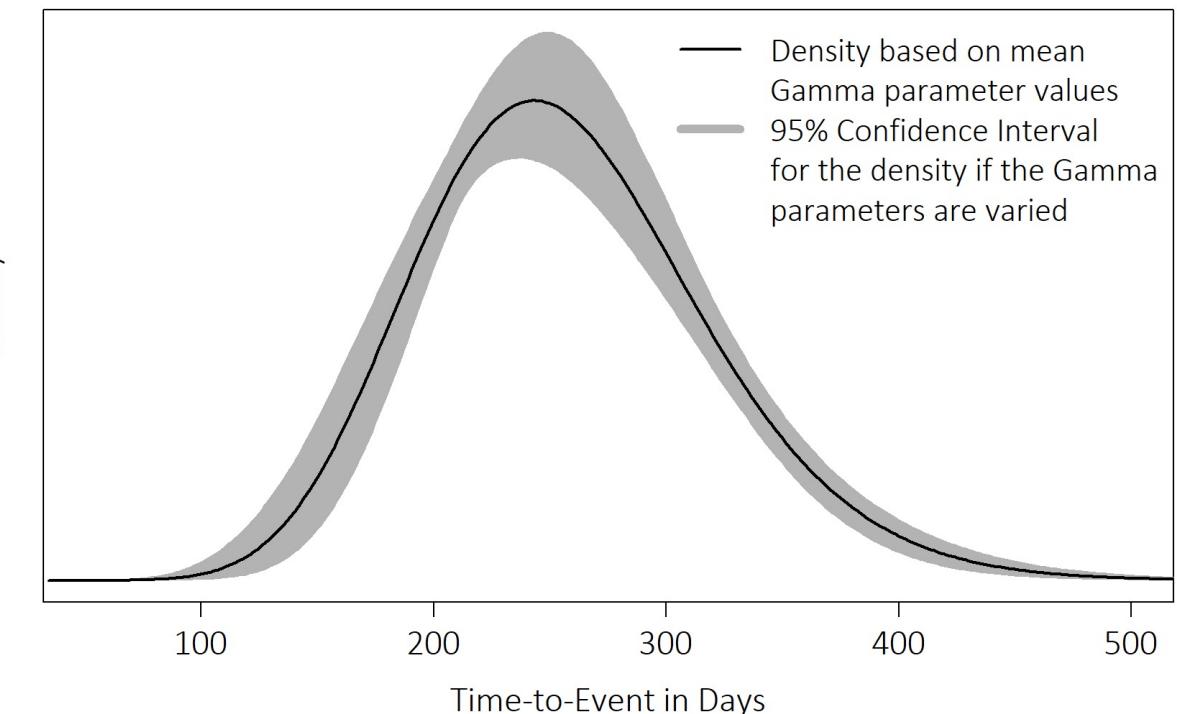
# Parameter uncertainty – DES specific approach

Illustration of outcomes: uncertainty in parameters of a Gamma distribution and resulting uncertainty in the time-to-event distribution

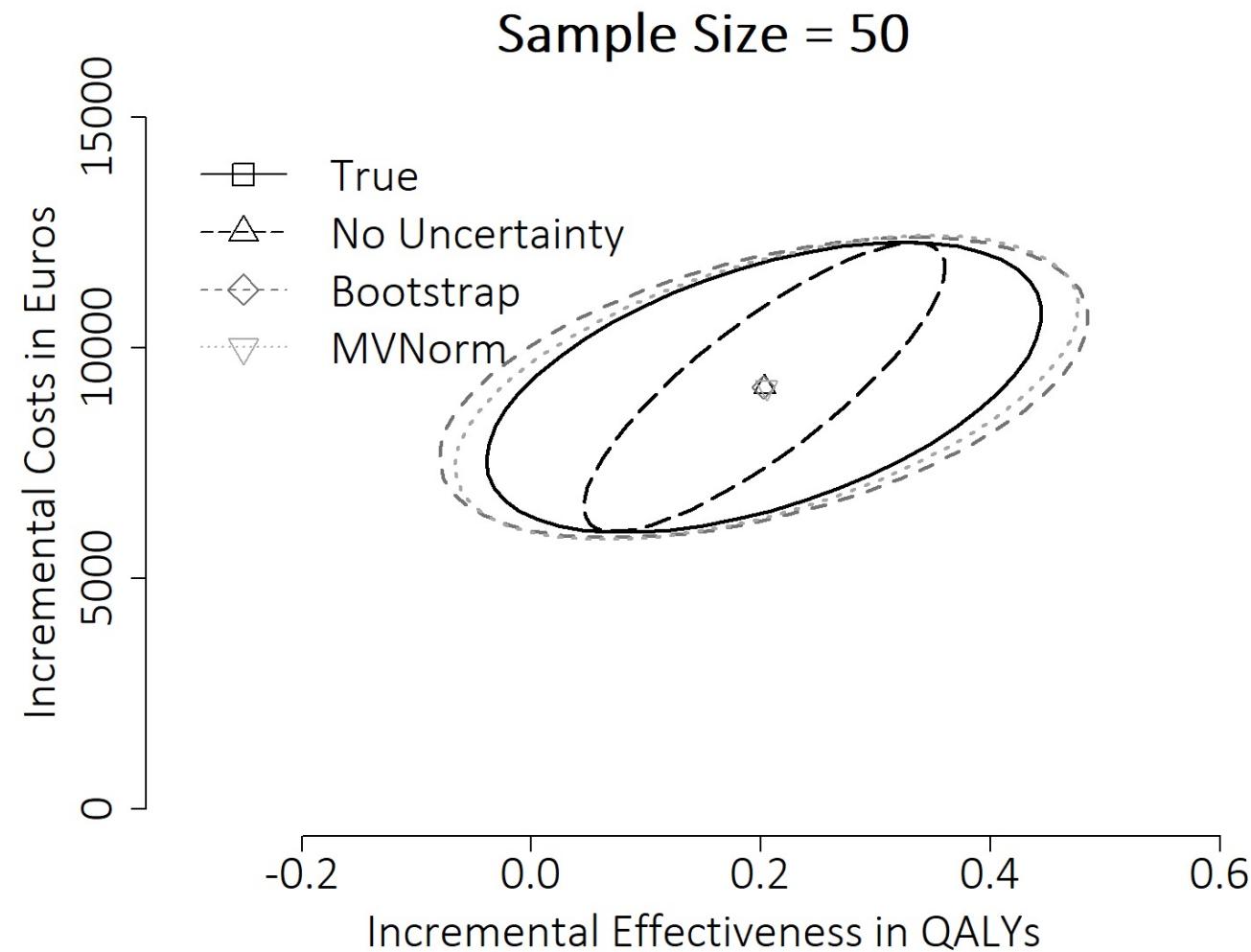
a) Gamma Distribution Parameter Estimates



b) Simulated Gamma Distribution Density Curves



# Parameter uncertainty – DES specific approach



Small sample sizes, of less than 100, 50, or even 20 individuals are more common than you think!

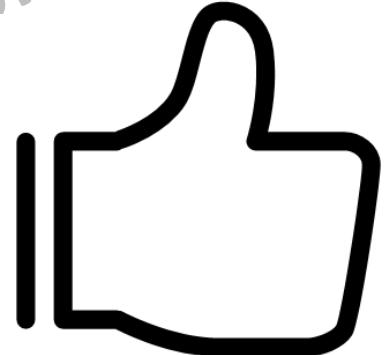
Following stratification for competing risks or when applying multivariable distributions

# Comparison of the 2 approaches

21JRSE

## Non-parametric bootstrapping

- Does not require any assumptions regarding underlying distributions
- Conserves the correlation between *all distributions* based on the IPD
- Less likely to result in extreme time-to-event values
- Small datasets or low number of events may produce extreme time-to-event values



## Using multivariate Normal distributions

- Assumes normality of distributions' parameter estimates (unlikely in small samples)
- Does not conserve the correlation between *all distributions* based on the IPD
- More often produces extreme time-to-event values

# Summary

**Uncertainty is important - all types**

In DES

- Always check for residual stochastic uncertainty (stable outcomes)
- Address uncertainty in parameters estimated for parametric probability distributions (preferably through bootstrapping)

DISCRETE EVENT SIMULATION 1 SMDM 2021 SHORT COURSE  
H. KOFFIJBERG, K. DECERIĆ, M. VAN DE VEN