Microsimulation modeling in R

Petros Pechlivanoglou, PhD
Zoe Kao, MA
Eline Krijkamp, MSc

41st Annual North American Meeting of the Society for Medical Decision Making

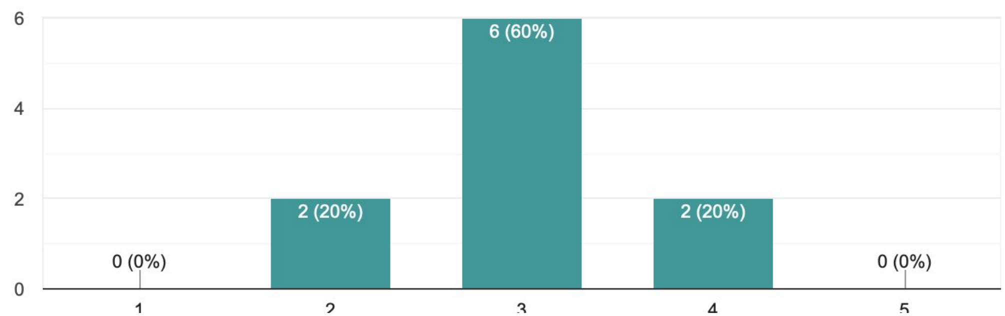
Portland, October 2019

Student introduction

Image: freepik.com

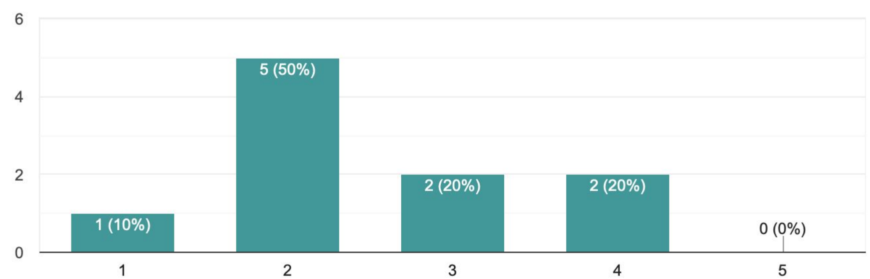
How would you rate your programming language skills?

10 responses



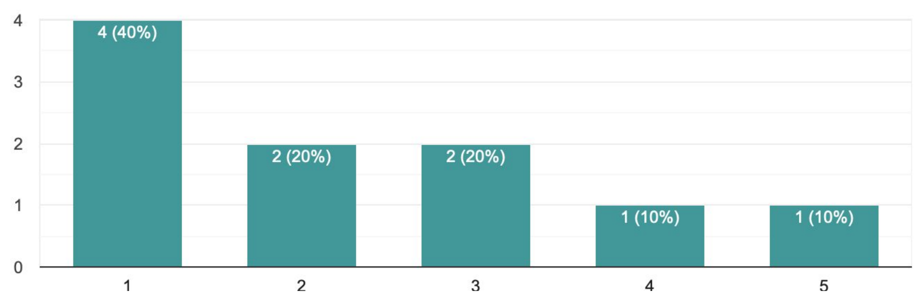
How your you rate your theoretical knowledge about microsimulation models?

10 responses



What is your experience with building decision models in programming languages?

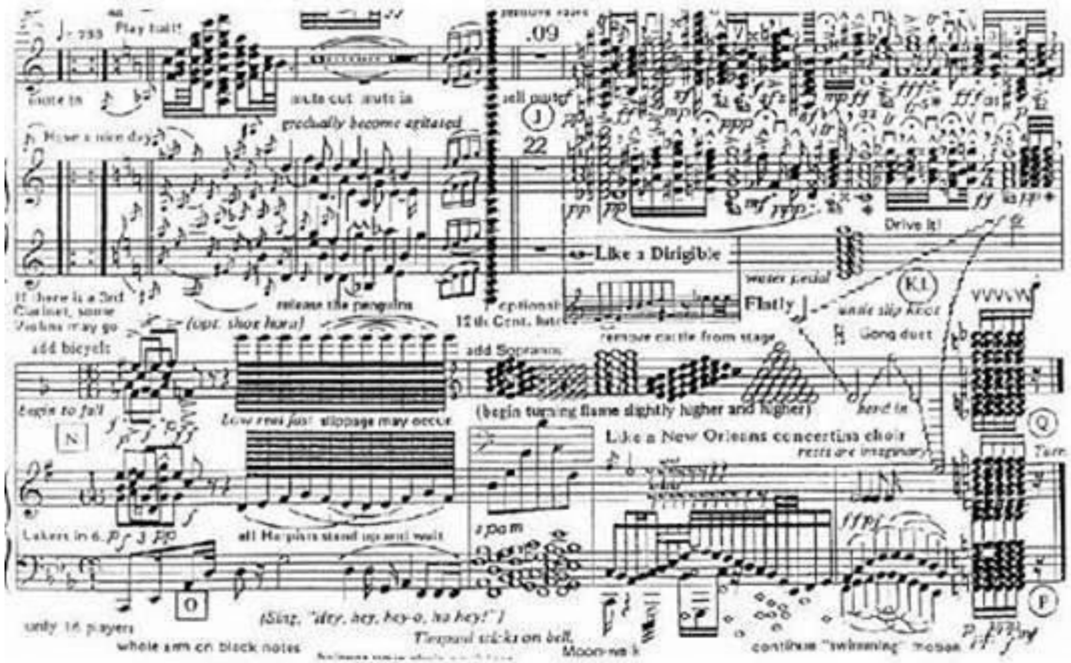
10 responses



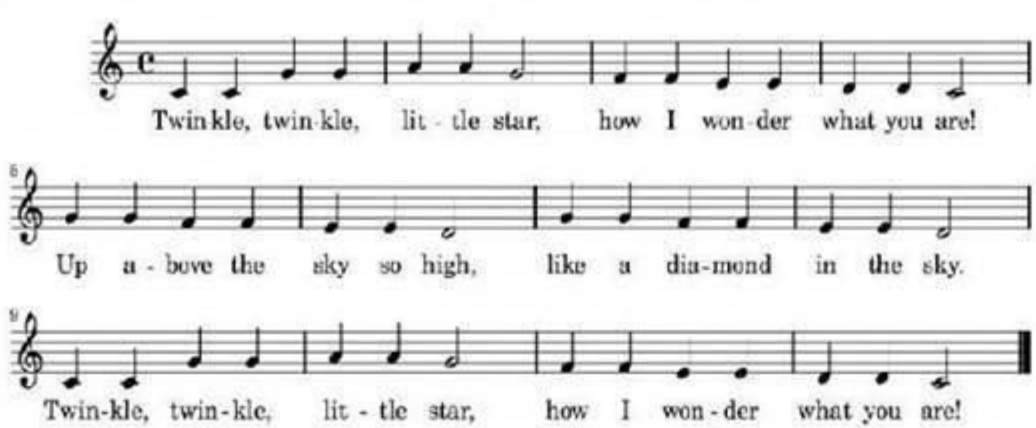
Survey results

1 is unexperienced 5 is expert

HOW THE SHEET LOOKS WHEN YOU'RE SIGHT READING:



HOW IT LOOKS AFTER YOU'VE PRACTICED:



Today

Part 1: Introduction + Model building

- Construct microsimulation models
- Visualize and analyze outputs
- Understand computational efficiency considerations in implementing a microsimulation

Coffee break

Part 2: State-residency and PSA

- Appreciate the advantages and challenges of using R in decision modeling

The DARTH Workgroup

- Materials for this workshop were largely developed by the Decision Analysis in R for Technologies in Health (DARTH) Workgroup

- Goals: To expand knowledge in decision analysis using R and develop educational materials to empower people to construct R-based decision models.

For more information

www.darthworkgroup.com

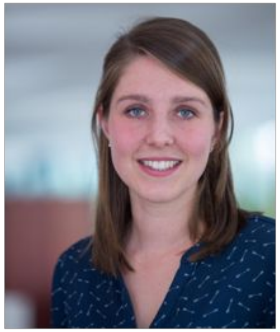
Tweet: @DARTHworkgroup

The DARTH workgroup

F. Alarid-Escudero PhD

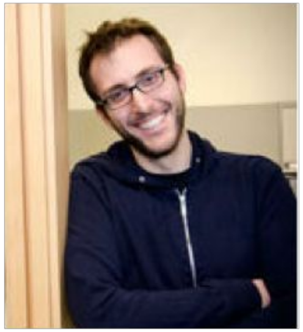
H. Jalal MD PhD

E. Enns PhD

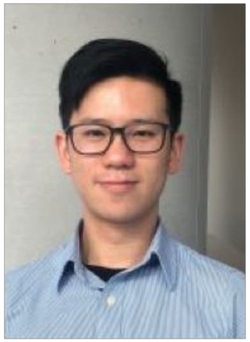


E. Krijkamp PhDc

M. Hunink MD



P. Pechlivanoglou PhD



Alan Yang MSc

Attribution and Acknowledgement

- R code provided with this workshop are yours to reuse and modify

```
################################################################################
# Please cite our publications when using this code
# - Jalal H, Pechlivanoglou P, Krijkamp E, Alarid-Escudero F, Enns E, Hunink MG.
# An Overview of R in Health Decision Sciences. Med Decis Making. 2017; 37(3): 735-746.
# - Krijkamp EM, Alarid-Escudero F, Enns EA, Jalal HJ, Hunink MGM, Pechlivanoglou P.
# Microsimulation modeling for health decision sciences using R: A tutorial.
# Med Decis Making. 2018;38(3):400-22.

################################################################################
```

Acknowledgement and citation information in code headers

darthpack

PharmacoEconomics
https://doi.org/10.1007/s40273-019-00837-x

PRACTICAL APPLICATION

Check for updates

A Need for Change! A Coding Framework for Improving Transparency in Decision Modeling

Fernando Alarid-Escudero[1] · Eline M. Krijkamp[2] · Petros Pechlivanoglou[3] · Hawre Jalal[4] · Szu-Yu Zoe Kao[5] · Alan Yang[6] · Eva A. Enns[5]

Prefix	Data type	Prefix	Variable type	Prefix	Variable type
<> (no prefix)	scalar	n	number	ly	life years
v	vector	p	probability	q	QALYs
m	matrix	r	rate	se	standard error
a	array	u	utility		
df	data frame	c	cost		
dtb	data table	hr	hazard ratio		
l	list	rr	relative risk		

Abstract

Monday, October 21, 2019

5:45 PM - 6:00 PM

DoubleTree by Hilton - 3 Sisters/Mt. Bachelor, 2nd Level

Introduction to discrete time microsimulation

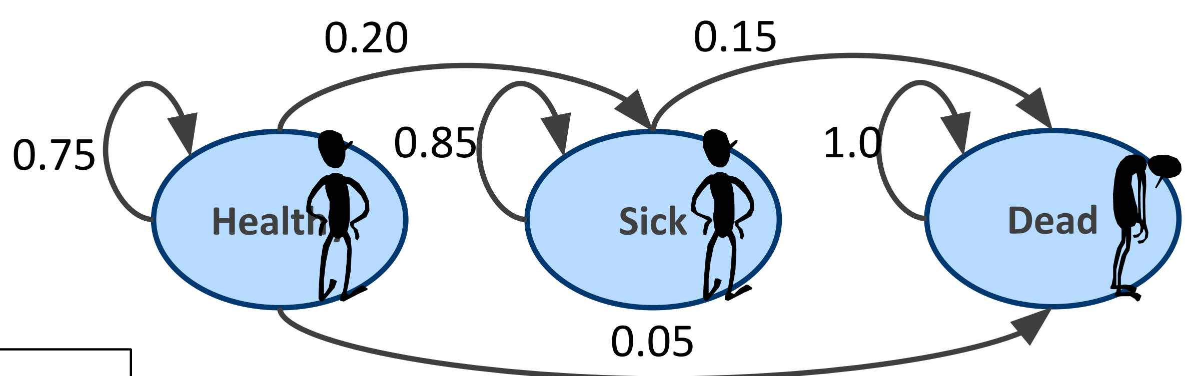
We thank Dr. Eva Enns from the DARTH workgroup for the basis of these microsimulation slides

What is microsimulation?

- Micro = individual-level model
- Simulation = stochastic implementation of a dynamic process
    - Reflects events experienced by an individual

Individual state-transition model

- Sometimes called "Markov Monte Carlo" or "First-Order Monte Carlo" or "Individual state transition model"
- Simulates *individual* disease progression through a state-transition model
  - Track individual's health state over time (can only be in one state at any given time)



t = 5:

General Microsimulation

- Need not explicitly follow a Markov model structure
- Track current state of individual as well as relevant history/characteristics
  - Need not be discrete categories; continuous measures possible
- Probabilities of simulated events can depend on
  - Individual characteristics (age, gender, etc.)
  - Full clinical history, time since clinical events

Pros/Cons of Microsimulation

Advantages

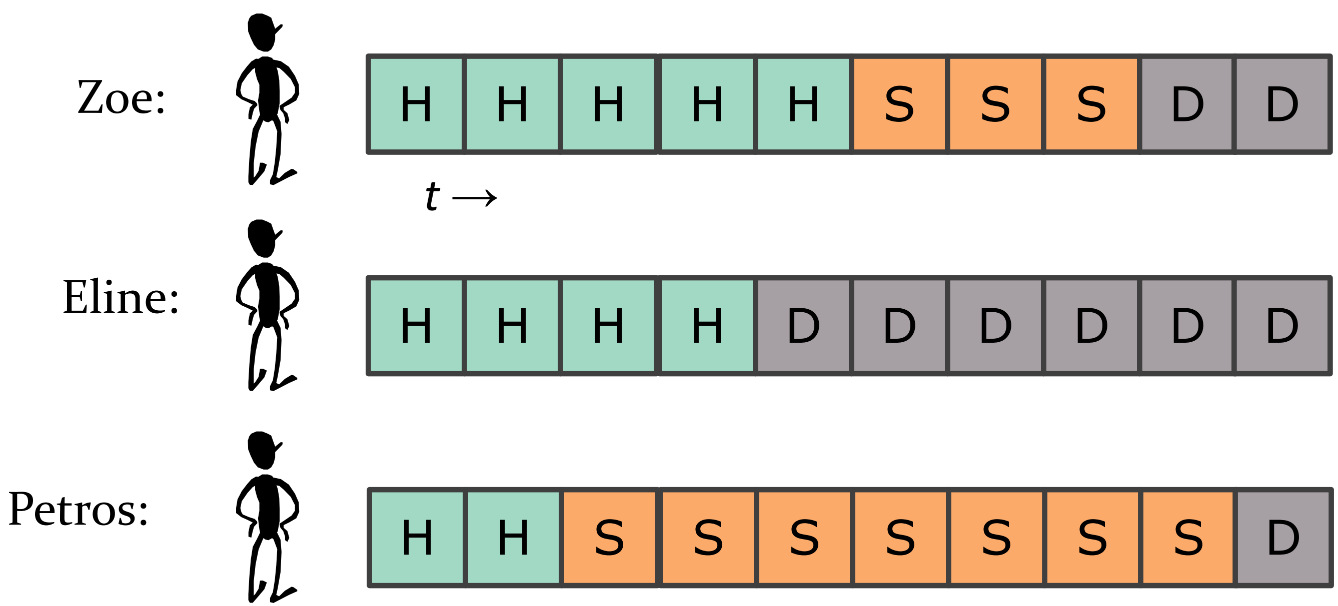
- Flexible model structure
- Easy to include:
  - Individual heterogeneity
  - Complex history-dependencies
  - Continuous health measures
  - Relation among individuals (network)

Disadvantages

- Complex to implement
- Computationally intensive
- Requires more data to inform model parameter values

Microsimulation Basics

- Simulate disease progression and health outcomes in an individual
- Simulate many individuals to estimate expected value and standard deviation of health outcomes over a large population



Efficiency in Microsimulation

- Microsimulation can be computationally intensive
- Simulating one individual at a time is an intuitive, but inefficient, approach

```r
for(i in 1:n_i) { # open individual loop
  for (t in 1:n_t) { # open time loop

     # simulation code here #

  } # close time loop
} # close individual loop
```

Tutorial

MDM
Medical Decision Making

Medical Decision Making
2018, Vol. 38(3) 400–422

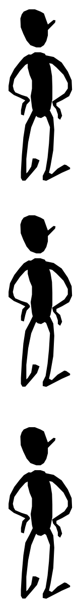
Reprints and permissions:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/0272989X18754513
journals.sagepub.com/home/mdm
SAGE

Microsimulation Modeling for Health Decision Sciences Using R: A Tutorial

Eline M. Krijkamp, Fernando Alarid-Escudero, Eva A. Enns, Hawre J. Jalal, M. G. Myriam Hunink, and Petros Pechlivanoglou

Efficiency in Microsimulation

- "Batch process" individual at each time step (one for-loops)



	t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8	t_9	t_{10}
Zoe:	H	H	H	H	H	S	S	S	D	D
Eline:	H	H	H	H	D	D	D	D	D	D
Petros:	H	H	S	S	S	S	S	S	S	D

iterative vs vectorized

Sample size	Time to run (in seconds)	
	sample()	samplev()
1,000	5.42	0.16
10,000	38.41	1.21
100,000	378.76	11.71
1,000,000	4538.80	128.79

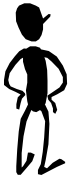
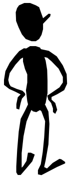
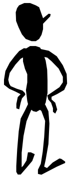
Microsimulation in R

- Generate a representative, virtual population
  - Sample characteristics from demographic data
  - Age distribution, M:F ratio, etc.

- Simulate the occurrence of events
  - Write functions that calculate individual-specific probabilities of different events
  - $p_{event1} = f(age, sex, health\ status, time\ since\ event, \ldots)$
  - Simulate events (and their consequences) over time using random numbers

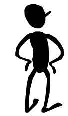
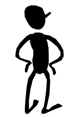
- Calculate population-level outcomes by averaging individual outcomes

Individuals Characteristics

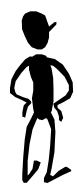
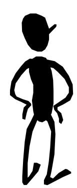
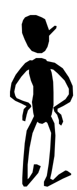
		age	sex	height	Country of birth
Zoe:		35	Female	1.55	Taiwan
Eline:		28	Female	1.68	NL
Petros		36	Male	1.89	Greece

Individuals Characteristics

t = 1

Zoe:		H	H	H	H	H	S	S	S	D	D
Eline:		H	H	H	H	D	D	D	D	D	D
Petros:		H	H	S	S	S	S	S	S	S	D

$t \rightarrow$

		age	sex	height	Country of birth	time Sick	p_HS
Zoe:		35	Female	1.55	Taiwan	0	0.3
Eline:		28	Female	1.68	NL	0	0.3
Petros		36	Male	1.89	Greece	0	0.4

Individuals Characteristics

t = 2

Zoe:	H	H	H	H	H	S	S	S	D	D
Eline:	H	H	H	H	D	D	D	D	D	D
Petros:	H	H	S	S	S	S	S	S	S	D

$t \rightarrow$

```
df_X$Age <- df_X$Age + 1
```

	age	sex	height	Country of birth	time Sick	p_HS
Zoe:	36	Female	1.55	Taiwan	0	0.3
Eline:	29	Female	1.68	NL	0	0.3
Petros	37	Male	1.89	Greece	0	0.4

Individuals Characteristics

t = 3

Zoe:	H	H	H	H	H	S	S	S	D	D
Eline:	H	H	H	H	D	D	D	D	D	D
Petros:	H	H	S	S	S	S	S	S	S	D

$t \rightarrow$

```
p_HD<- df_X$p_HD[m_M[, t] == "H"]
```

	age	sex	height	Country of birth	time Sick	p_HS
Zoe:	37	Female	1.55	Taiwan	0	0.3
Eline:	30	Female	1.68	NL	0	0.3
Petros	38	Male	1.89	Greece	1	--

Individuals Characteristics

Zoe:	H	H	H	H	H	S	S	S	D	D
Eline:	H	H	H	H	D	D	D	D	D	D
Petros:	H	H	S	S	S	S	S	S	S	D

$t \rightarrow$

t = 4

	age	sex	height	Country of birth	time Sick	p_HS
Zoe:	38	Female	1.55	Taiwan	0	0.3
Eline:	31	Female	1.68	NL	0	0.3
Petros	39	Male	1.89	Greece	2	--

Individuals Characteristics

t = 5

Zoe:	H	H	H	H	H	S	S	S	D	D
Eline:	H	H	H	H	D	D	D	D	D	D
Petros:	H	H	S	S	S	S	S	S	S	D

$t \rightarrow$

	age	sex	height	Country of birth	time Sick	p_HS
Zoe:	39	Female	1.55	Taiwan	0	0.3
Eline:	32	Female	1.68	NL	0	--
Petros	40	Male	1.89	Greece	3	--

Individuals Characteristics

t = 6

Zoe:	H	H	H	H	H	S	S	S	D	D
Eline:	H	H	H	H	D	D	D	D	D	D
Petros:	H	H	S	S	S	S	S	S	S	D

$t \rightarrow$

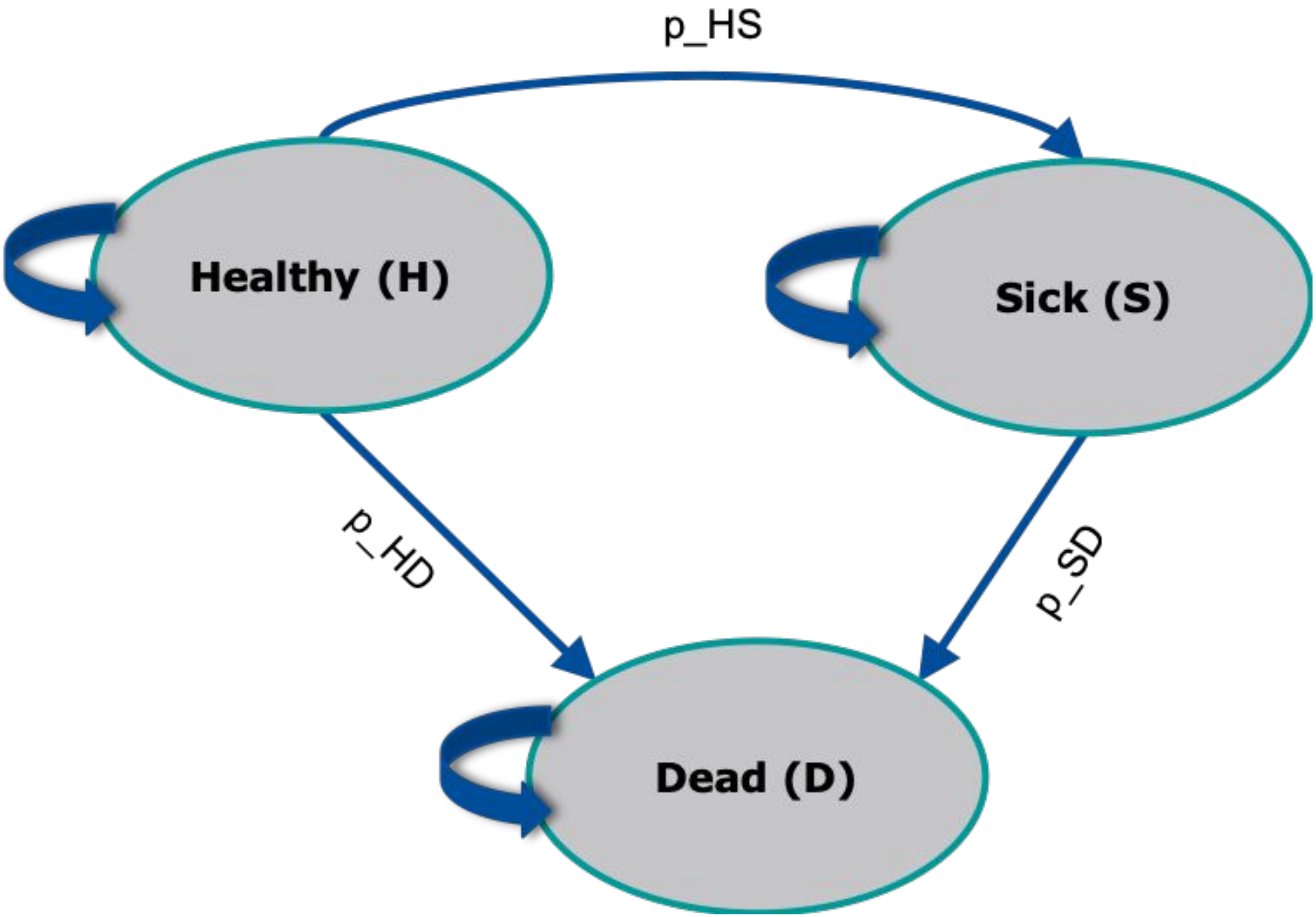
	age	sex	height	Country of birth	time Sick	p_HS
Zoe:	40	Female	1.55	Taiwan	1	--
Eline:	33	Female	1.68	NL	0	--
Petros	41	Male	1.89	Greece	4	--

R session

3-state example

- Three-state model of disease: Healthy, Sick, Dead
- Simulate a population of 10,000 individuals
- Individual characteristics
  - Sex – assume equal proportion of women and men
- p_HD is sex-dependent

Female	0.0382
Male	0.0463


p_HS
Healthy (H)
Sick (S)
p_HD
p_SD
Dead (D)

Functions

```
calculateMean <- function (x){
  mean <- sum(x)/length(x)
}
```

Structure of our code

Specify all the input parameters

- transition probabilities, cycle length etc

Generate sample with individual (baseline) characteristics X

- age, sex etc

Specify functions

- $Probs(m, x)$
- $Costs\ (m, x)$
- $Effs\ \ (m, x)$

$MicroSim(\,)$

$C_0 = Costs(M_0, X_0)$

$E_0 = Effs(M_0, X_0)$

for $t = 1$ to nt do

$p = Probs(M_t, X_t)$

$M_{t+1} \sim samplev\ (n, p)$

$Update\ X_{t+1}$

$C_{t+1} = Costs(M_{t+1}, X_{t+1})$

$E_{t+1} = Effs\ (M_{t+1}, X_{t+1})$

End $MicroSim$

Run $MicroSim$

Microsimulation exercise

Sick – Sicker model

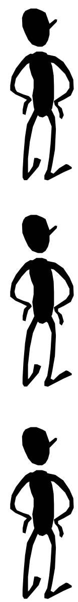
See exercise instructions

State-residence

State-residency

- Probability might be dependent on how long someone is in a state
  - The probability of recovery depends on the duration of being sick

Duration of being sick (step)	p_SH
1	0.9
2	0.5
3+	0.1



	t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8	t_9	t_{10}
Zoe:	H	S	S	S	H	H	H	H	H	H
Eline:	H	H	H	S	S	H	H	H	H	H
Petros:	H	H	S	S	S	S	S	S	S	H

Individuals Characteristics

t = 1

Zoe:	H	S	S	S	H	H	H	H	H	H
Eline:	H	H	H	S	S	H	H	H	H	H
Petros:	H	H	S	S	S	S	S	S	S	H

$t \rightarrow$

	time Sick	p_SH
Zoe:	0	--
Eline:	0	--
Petros	0	--

Individuals Characteristics

t = 2

Zoe:	H	S	S	S	H	H	H	H	H	H
Eline:	H	H	H	S	S	H	H	H	H	H
Petros:	H	H	S	S	S	S	S	S	S	H

$t \rightarrow$

	time Sick	p_SH
Zoe:	1	0.9
Eline:	0	--
Petros	0	--

Individuals Characteristics

t = 3

Zoe:	H	S	S	S	H	H	H	H	H	H
Eline:	H	H	H	S	S	H	H	H	H	H
Petros:	H	H	S	S	S	S	S	S	S	H

$t \rightarrow$

	time Sick	p_SH
Zoe:	2	0.5
Eline:	0	--
Petros	1	0.9

Individuals Characteristics

Zoe:	H	S	S	S	H	H	H	H	H	H
Eline:	H	H	H	S	S	H	H	H	H	H
Petros:	H	H	S	S	S	S	S	S	S	H

$t \rightarrow$

$t = 4$

	time Sick	p_SH
Zoe:	3	0.1
Eline:	1	0.9
Petros	2	0.5

Individuals Characteristics

$t = 5$

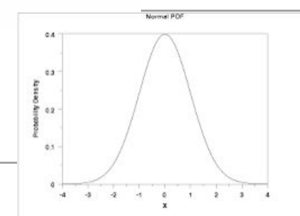
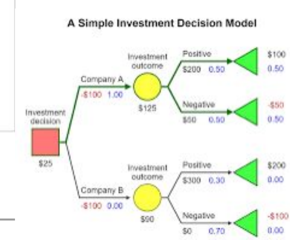
Zoe:	H	S	S	S	H	H	H	H	H	H
Eline:	H	H	H	S	S	H	H	H	H	H
Petros:	H	H	S	S	S	S	S	S	S	H

$t \rightarrow$

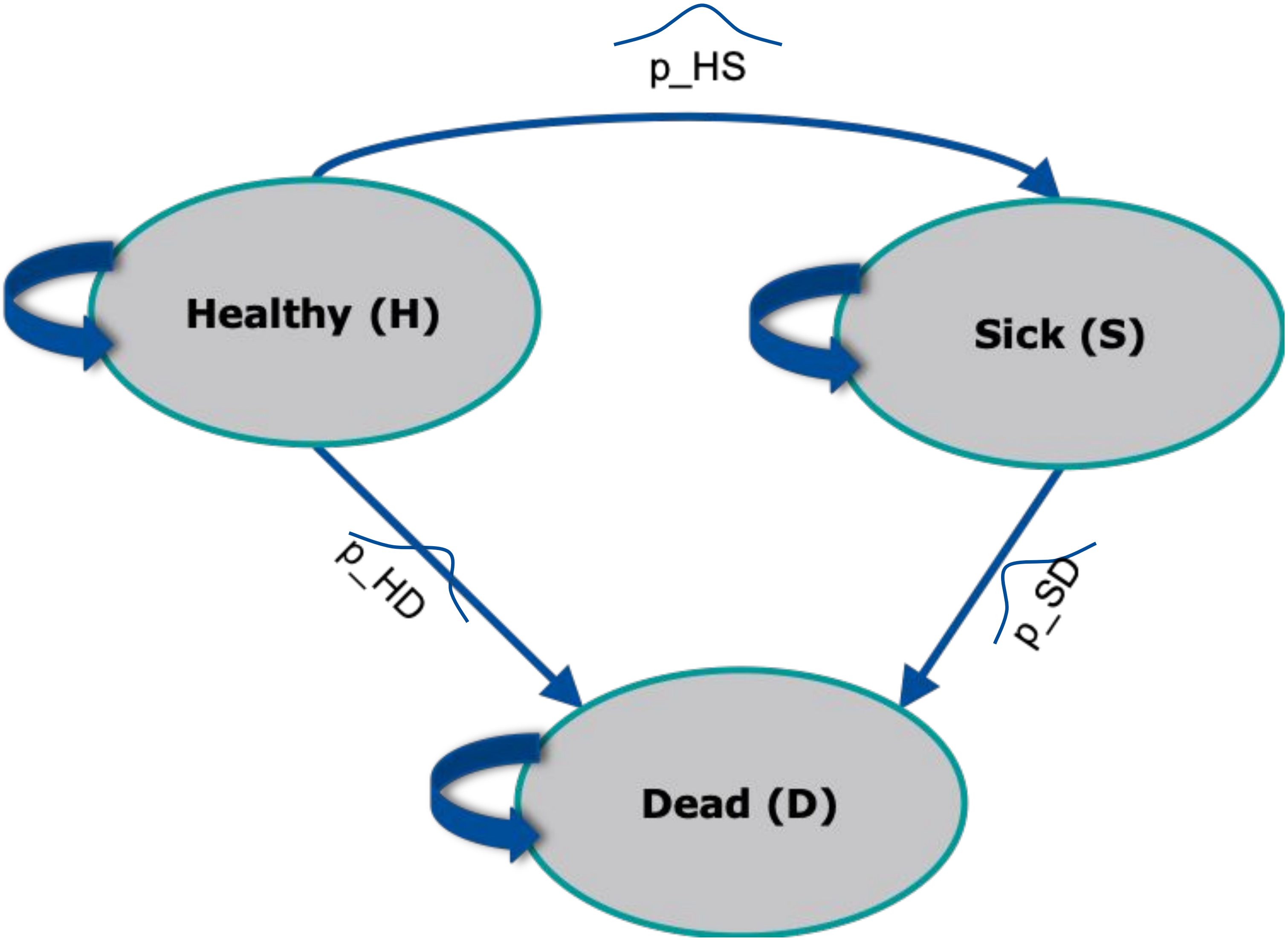
	time Sick	p_SH
Zoe:	0	--
Eline:	2	0.5
Petros	3	0.1

PSA

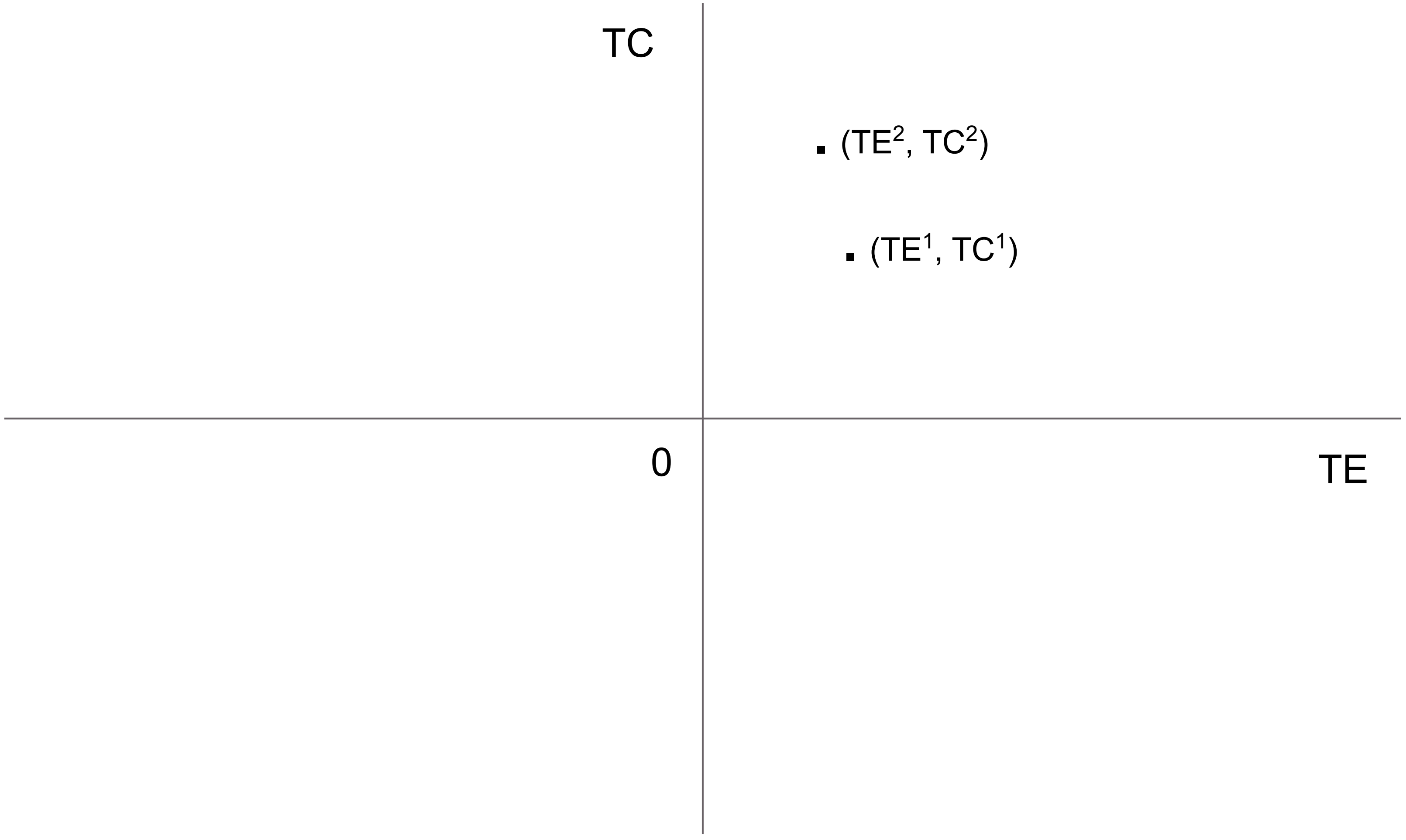
Heterogeneity vs. uncertainty

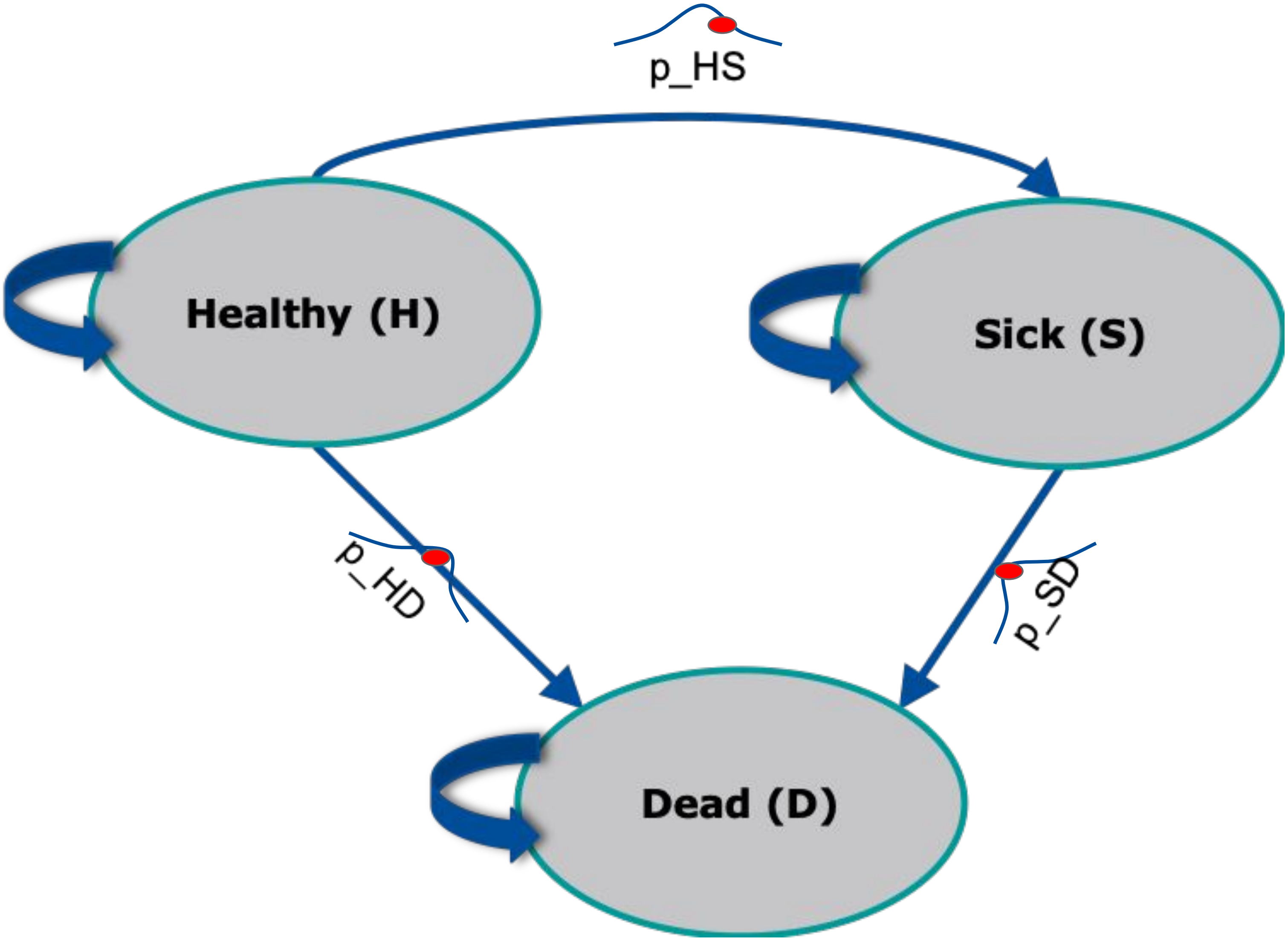
Heterogeneity	1st order uncertainty
Probability/utility/costs depent on	- age, sex, risk factors - events, treatment - duration in a state
Uncertainty	**2nd order uncertainty**
Stochastic uncertainty	
Parameter uncertainty	
Model uncertainty	

Slides based on RDS288 course Prof. Myriam Hunink

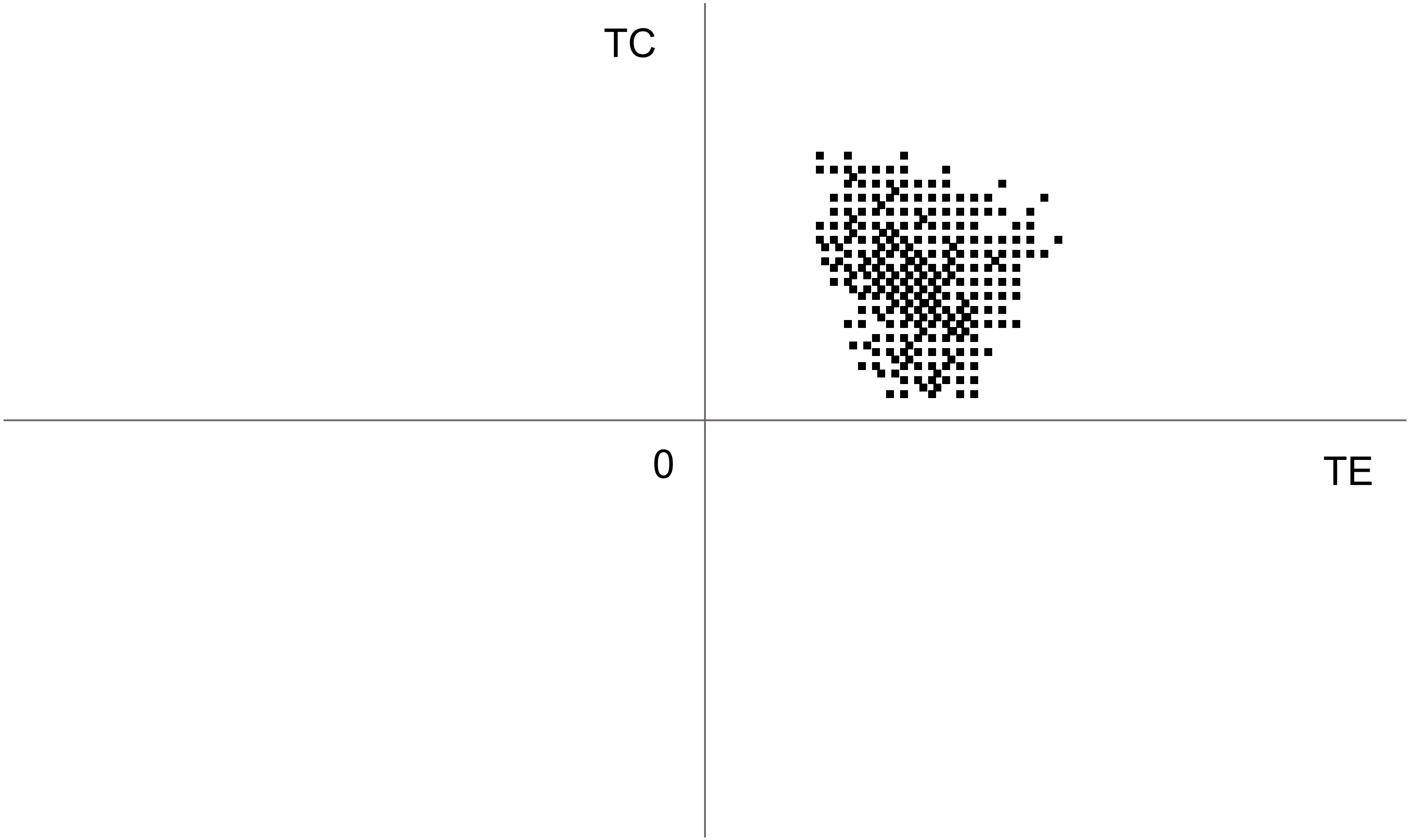

p_HS
Healthy (H)
Sick (S)
p_HD
p_SD
Dead (D)

Presenting the PSA results




p_HS
Healthy (H)
Sick (S)
p_HD
p_SD
Dead (D)

Presenting the PSA results



PSA in R

Common structure among (most of) distributions in R

- "q"+ *dist.* (e.g. `qnorm()`): quantile function
- "d"+ *dist.*(e.g. `dnorm()`): density function
- **"p"+ *dist.***(e.g. `pnorm()`)**: c.d.f function**
- **"r" + *dist.***(e.g. `rnorm()`)**: random number generating function**

`sample()`:Random number sampling with(out) replacement and weights

- multinomial distribution
- bootstrapping (if a dataset of paranters is available)

DARTH Publications

An Overview of R in Health Decision Sciences

Hawre Jalal, MD, PhD, Petros Pechlivanoglou, MSc, PhD, Eline Krijkamp, MSc, Fernando Alarid-Escudero, MSc, Eva Enns, MS, PhD, M. G. Myriam Hunink, MD, PhD

Microsimulation Modeling for Health Decision Sciences Using R: A Tutorial

Eline M. Krijkamp, Fernando Alarid-Escudero, Eva A. Enns, Hawre J. Jalal, M. G. Myriam Hunink, and Petros Pechlivanoglou

MDM
Medical Decision Making
Medical Decision Making
2018, Vol. 38(3) 400–422
© The Author(s) 2018
Reprints and permissions:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/0272989X18754513
journals.sagepub.com/home/mdm
SAGE

PharmacoEconomics
https://doi.org/10.1007/s40273-019-00837-x

PRACTICAL APPLICATION

Check for updates

A Need for Change! A Coding Framework for Improving Transparency in Decision Modeling

Fernando Alarid-Escudero[1] · Eline M. Krijkamp[2] · Petros Pechlivanoglou[3] · Hawre Jalal[4] · Szu-Yu Zoe Kao[5] · Alan Yang[6] · Eva A. Enns[5]

New Results

Comment on this paper

A Multidimensional Array Representation of State-Transition Model Dynamics

Eline M. Krijkamp, Fernando Alarid-Escudero, Eva A. Enns, Petros Pechlivanoglou, M.G. Myriam Hunink, Hawre J. Jalal

doi: https://doi.org/10.1101/670612

Acknowledgement

GORDON AND BETTY MOORE FOUNDATION

SMDM
Society for Medical Decision Making
Better Health through Better Decisions.

Thank you

Follow us: @DARTHworkgroup