Microsimulation modeling in R
- Theory -

DARTH workgroup

info@DARTHworkgroup.com
@DARTHworkgroup

Introduction to discrete time microsimulation

The aim of this session:

- to give a short introduction to the theory of discrete time microsimulation models
- define the terminology we use

What is discrete time microsimulation?

- Micro = individual-level
- Simulation = imitation of a situation or process
- Discrete time = fixed time intervals
  - Reflects events experienced by an individual
  - Stochastic implementation of a dynamic process

Microsimulation terminology

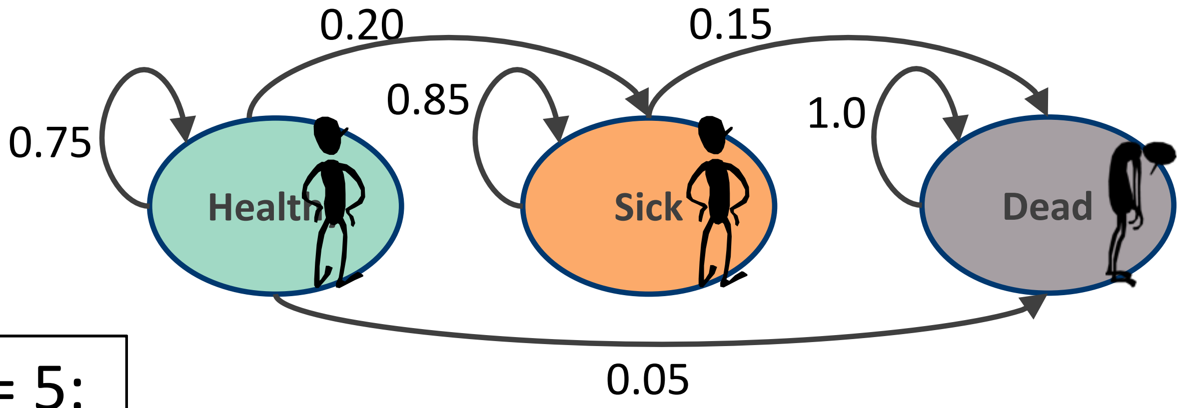
- Sometimes called “Markov Monte Carlo” or “First-Order Monte Carlo” or “**Individual state-transition model**”

- Need not explicitly follow a state-transition model structure

- In our courses we refer to discrete time individual state-transition models when we talk about a microsimulation model

Simple example

- 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

- 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 and Cons

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

http://darthworkgroup.com/

https://github.com/organizations/DARTH-git

https://www.linkedin.com/groups/8635339

Microsimulation modeling in R
- Efficiency -

DARTH workgroup

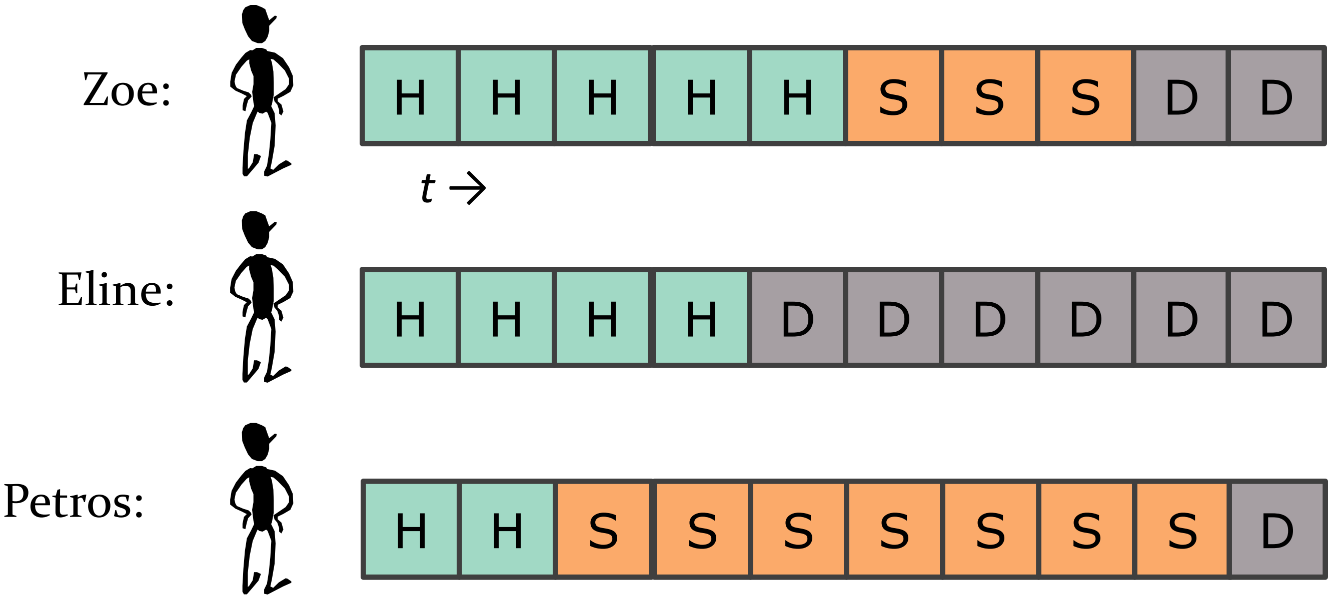
info@DARTHworkgroup.com
@DARTHworkgroup

The aim of this session is:

- to discuss the fundamental concept of microsimulation implementation
- to discuss batch-processing to improve efficiency

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**

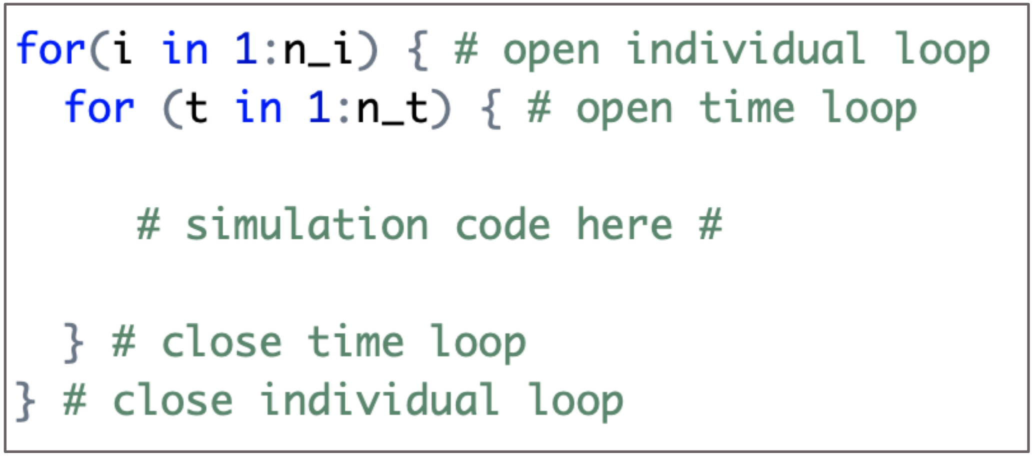


Zoe:	H	H	H	H	H	S	S	S	D	D
Eline:	H	H	H	H	D	D	D	D	D	D
Petros:	H	H	S	S	S	S	S	S	S	D

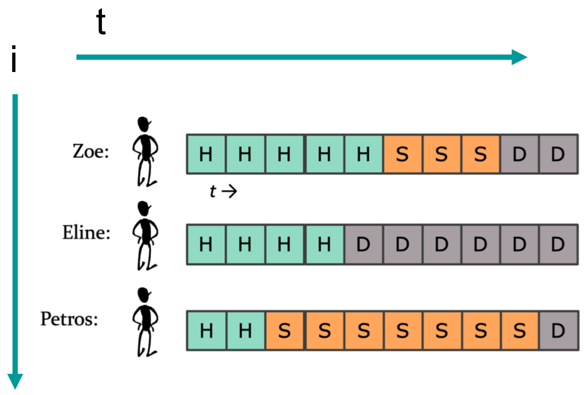
t →

Implementation

- Simulating one individual at a time is an intuitive approach…



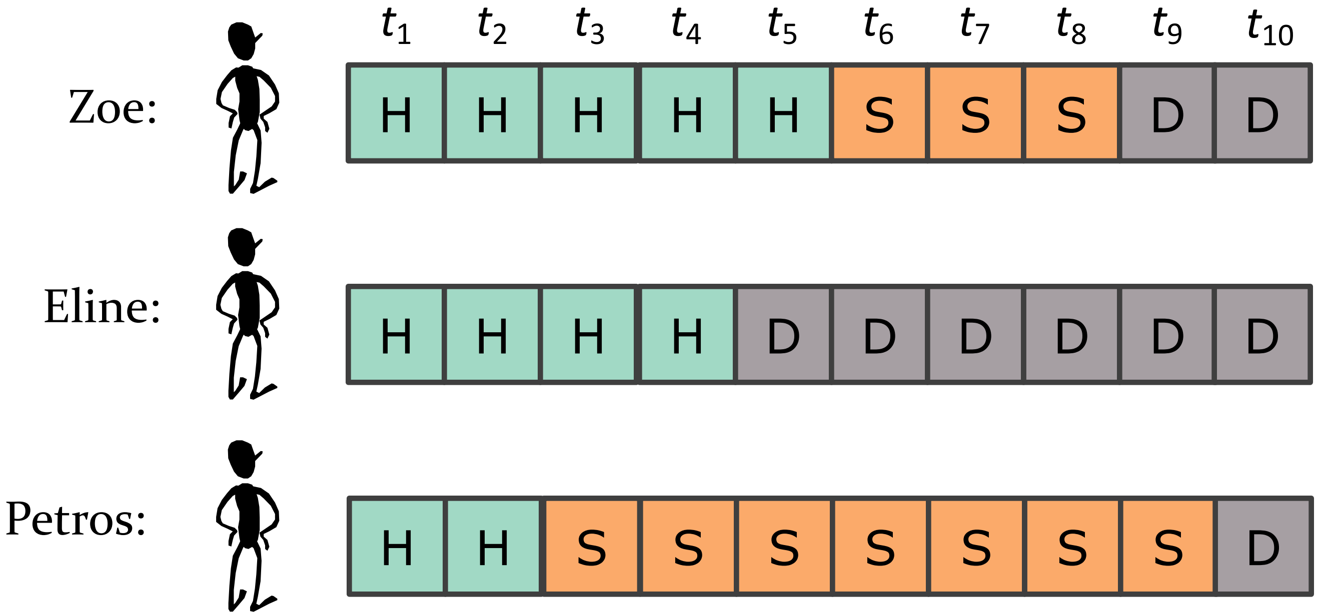
```r
for(i in 1:n_i) { # open individual loop
  for (t in 1:n_t) { # open time loop

     # simulation code here #

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



…. however, inefficent and can make a microsimulation computationally intensive

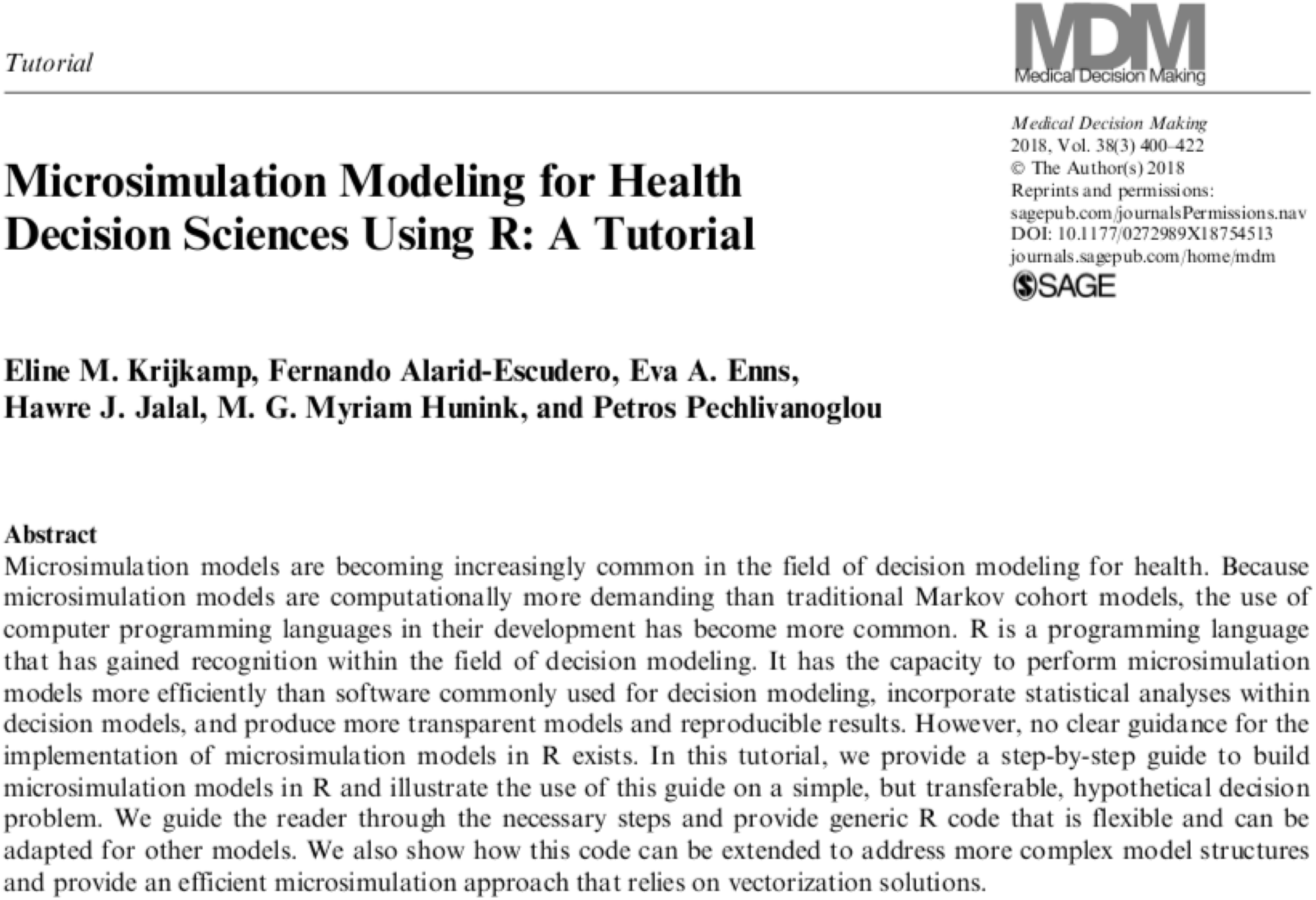
Efficiency in Microsimulation

- ”**Batch process**” individual at each time step (one for-loops, function `samplev`)



	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

Microsimulation Tutorial



Tutorial

Medical Decision Making

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

DOI: 10.1177/0272989X18754513
journals.sagepub.com/home/mdm
SAGE

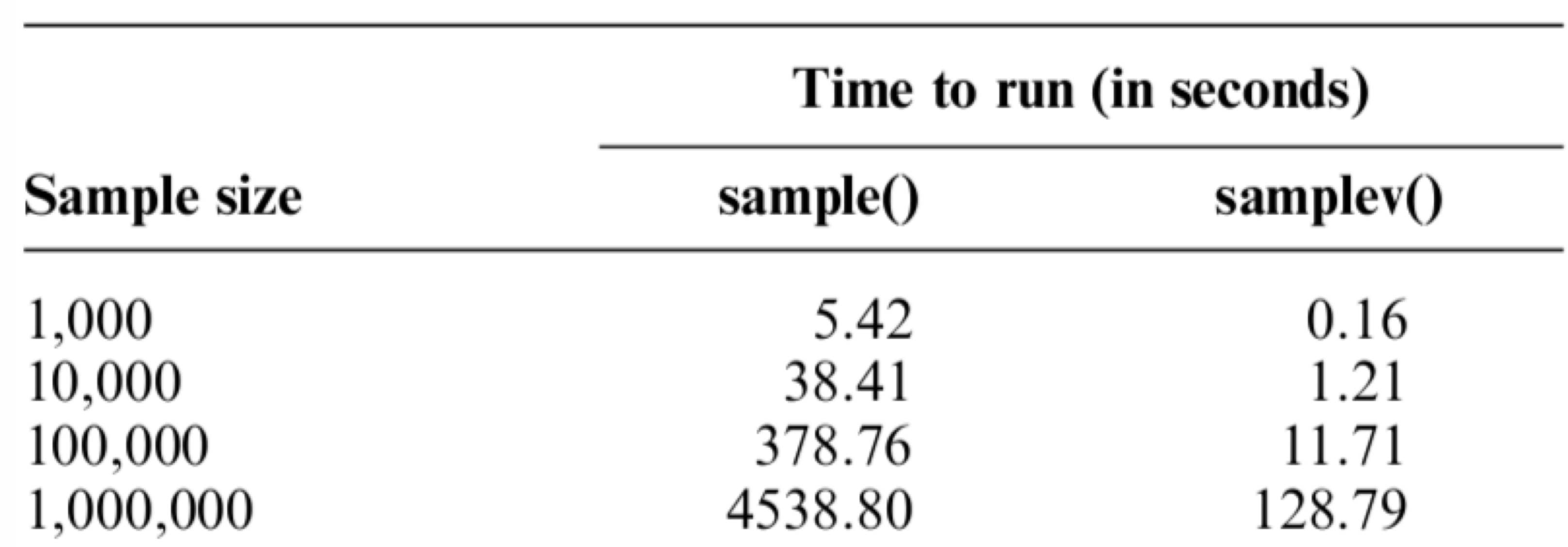
Microsimulation Modeling for Health Decision Sciences Using R: A Tutorial

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

Abstract

Microsimulation models are becoming increasingly common in the field of decision modeling for health. Because microsimulation models are computationally more demanding than traditional Markov cohort models, the use of computer programming languages in their development has become more common. R is a programming language that has gained recognition within the field of decision modeling. It has the capacity to perform microsimulation models more efficiently than software commonly used for decision modeling, incorporate statistical analyses within decision models, and produce more transparent models and reproducible results. However, no clear guidance for the implementation of microsimulation models in R exists. In this tutorial, we provide a step-by-step guide to build microsimulation models in R and illustrate the use of this guide on a simple, but transferable, hypothetical decision problem. We guide the reader through the necessary steps and provide generic R code that is flexible and can be adapted for other models. We also show how this code can be extended to address more complex model structures and provide an efficient microsimulation approach that relies on vectorization solutions.

iterative vs vectorized (“batch proces”)



	Time to run (in seconds)	
Sample size	**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

🌐 http://darthworkgroup.com/

GitHub https://github.com/organizations/DARTH-git

in https://www.linkedin.com/groups/8635339

Microsimulation modeling in R
- The implementation concept -

DARTH workgroup

info@DARTHworkgroup.com
@DARTHworkgroup

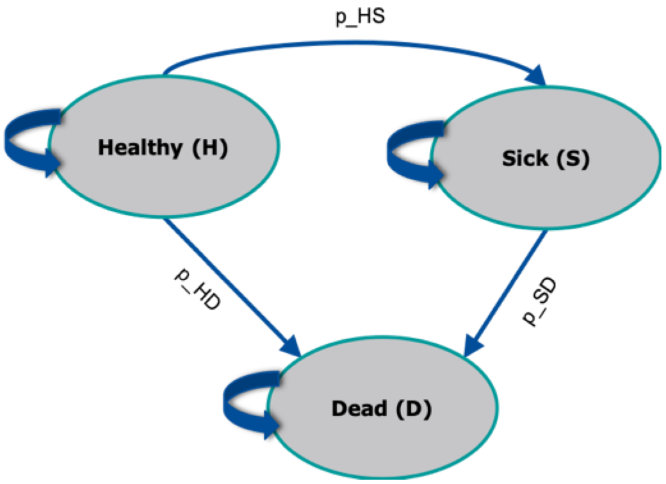
The aim of this session:

- to discuss the structures and steps that are needed for the implementation of a microsimulation in R
- Example with time constant probabilities
- Discuss time dependent probabilities

3-state example

- Three-state model of disease: Healthy, Sick, Dead
- Time horizon: 10 annual cycles
- Probability of transitioning from healthy to dead is sex-dependent

Female	0.0382
Male	0.0463



Individual characteristics

- Sex – assume equal proportion of women and men

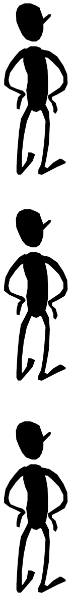
Microsimulation in R

- Generate a representative, virtual population (`df_X`)
  - Sample characteristics from demographic data
  - Male:Female ratio, age distribution, etc.

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

- Calculate population-level outcomes by averaging individual outcomes

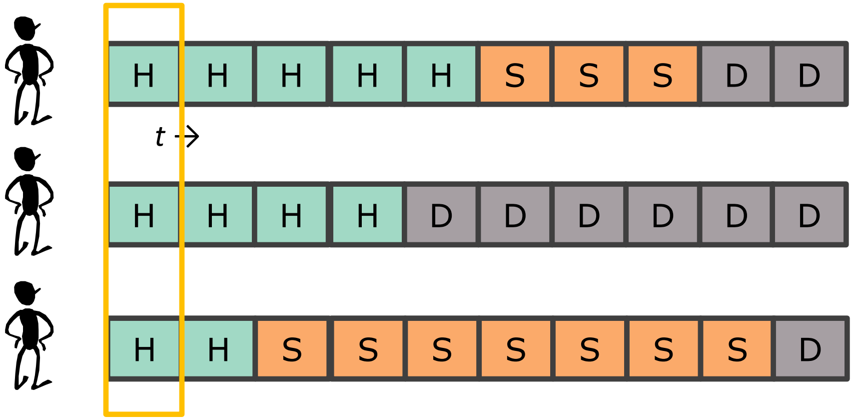
Structure with Individuals Characteristics (`df_X`)



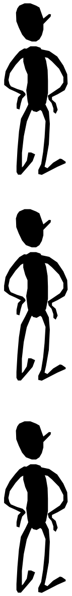
	age	sex	height	Country of birth
Zoe	35	Female	1.55	Taiwan
Eline	28	Female	1.68	NL
Petros	36	Male	1.89	Greece

Individuals Characteristics

t = 1

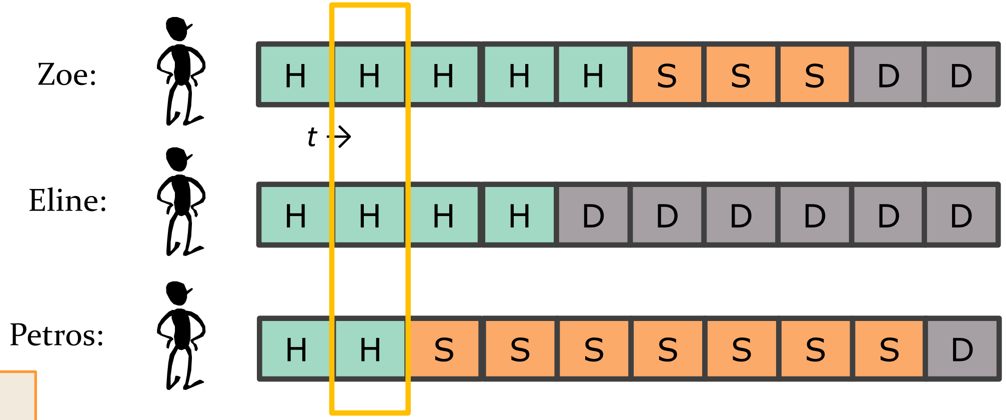


Zoe:	H	H	H	H	H	S	S	S	D	D
Eline:	H	H	H	H	D	D	D	D	D	D
Petros:	H	H	S	S	S	S	S	S	S	D



	age	sex	height	Country of birth	p_HD
Zoe	35	Female	1.55	Taiwan	0.3
Eline	28	Female	1.68	NL	0.3
Petros	36	Male	1.89	Greece	0.4

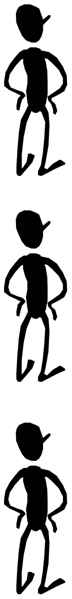
Update individuals Characteristics



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 = 2

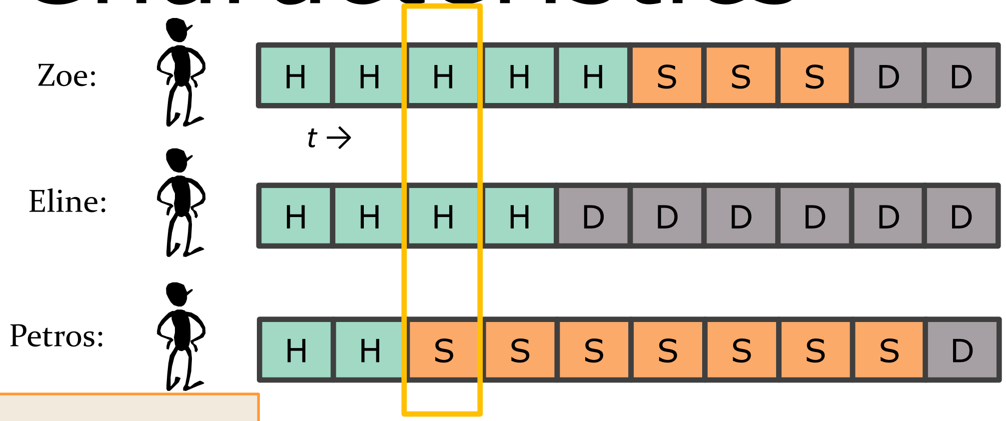
```
df_X$age <- df_X$age + 1
```



	age	sex	height	Country of birth	p_HD
Zoe	36	Female	1.55	Taiwan	0.3
Eline	29	Female	1.68	NL	0.3
Petros	37	Male	1.89	Greece	0.4

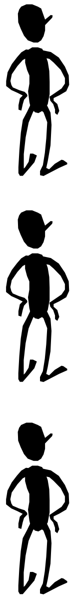
Individuals Characteristics

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

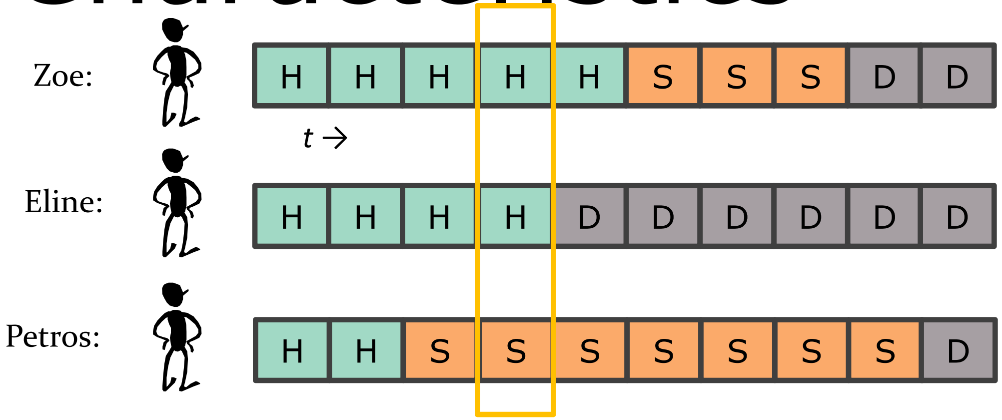
```
p_HD <- p_HD_all[M_t == "H", "p_HD"]
```



	age	sex	height	Country of birth	p_HD
Zoe	37	Female	1.55	Taiwan	0.3
Eline	30	Female	1.68	NL	0.3
Petros	38	Male	1.89	Greece	--

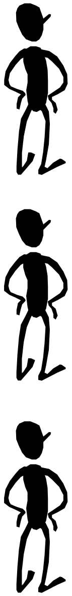
Individuals Characteristics

t = 4



t →

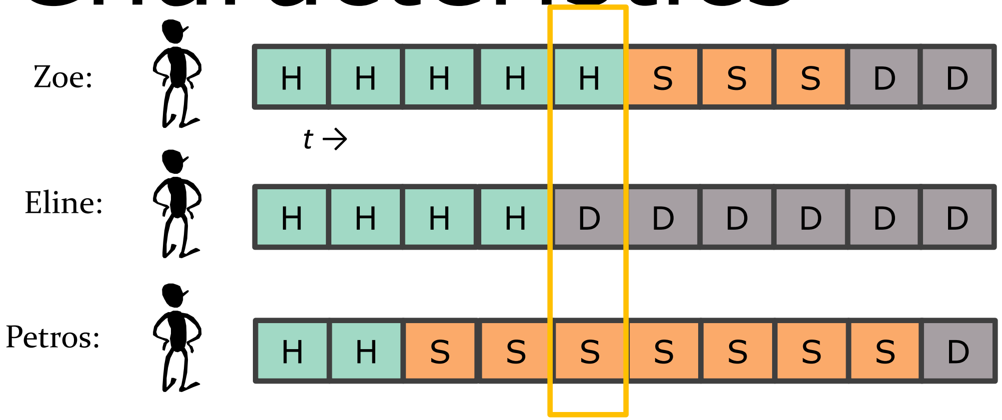
Zoe:	H	H	H	H	H	S	S	S	D	D
Eline:	H	H	H	H	D	D	D	D	D	D
Petros:	H	H	S	S	S	S	S	S	S	D



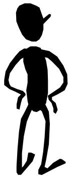
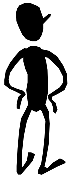
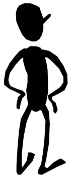
	age	sex	height	Country of birth	p_HD
Zoe	38	Female	1.55	Taiwan	0.3
Eline	31	Female	1.68	NL	0.3
Petros	39	Male	1.89	Greece	--

Individuals Characteristics

t = 5

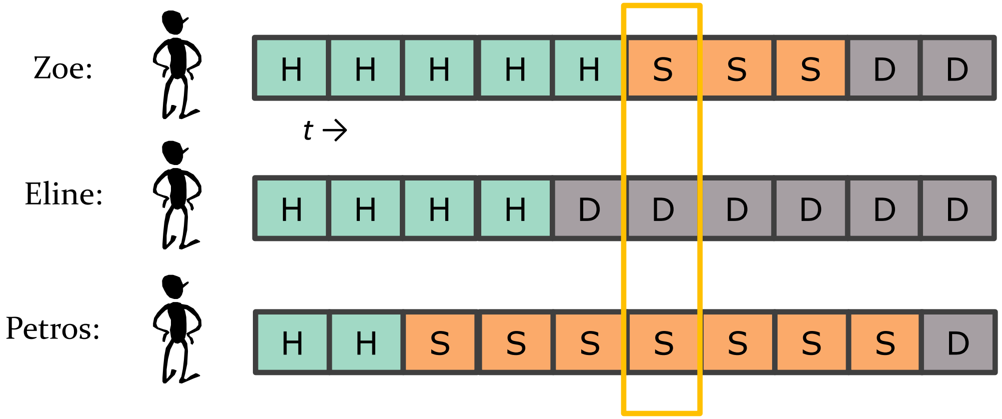


Zoe:	H	H	H	H	H	S	S	S	D	D
Eline:	H	H	H	H	D	D	D	D	D	D
Petros:	H	H	S	S	S	S	S	S	S	D

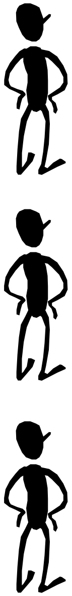
		age	sex	height	Country of birth	p_HD
Zoe		39	Female	1.55	Taiwan	0.3
Eline		32	Female	1.68	NL	--
Petros		40	Male	1.89	Greece	--

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



	age	sex	height	Country of birth	p_HD
Zoe	40	Female	1.55	Taiwan	--
Eline	33	Female	1.68	NL	--
Petros	41	Male	1.89	Greece	--

Time dependency

- Time constant probabilities in previous example
- Most models need time dependent probabilities
- The concept of indexing values based on personal characteristics works similar
    - age is updated in the `df_X` structure
    - update the age specific probability of dying
    - recalculate other parameters related to age

http://darthworkgroup.com/

https://github.com/organizations/DARTH-git

https://www.linkedin.com/groups/8635339

Microsimulation modeling in R
- State-residence -

DARTH workgroup

info@DARTHworkgroup.com
@DARTHworkgroup

The aim of this session:

- is to discuss state-residence in the context of implementation of a microsimulation in R

Time dependency

- **Since start of the model**
    - Transition probability often depend on age
        - Background mortality
        - Risk of developing disease or experiencing an event

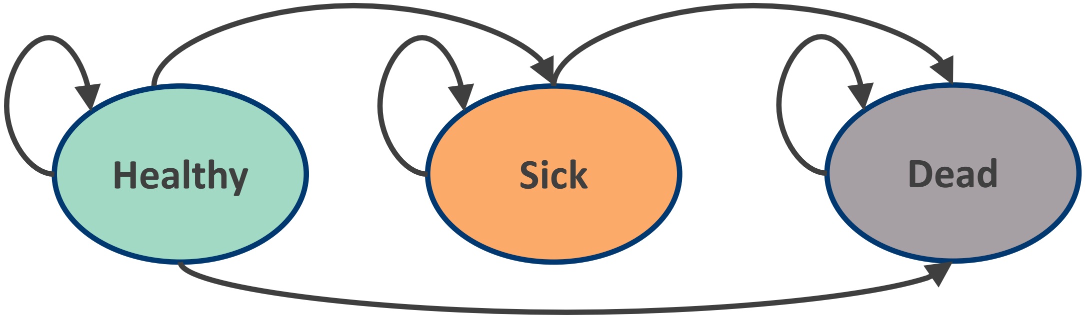
- **Depending on state residency**
    - Some transition probabilities depend on time since an event, not age
        - e.g., The risk of developing recurrence among newly diagnosed cancer patients declines with time

State-residency

- **Probability** might be dependent on how long someone is in a state

Example:

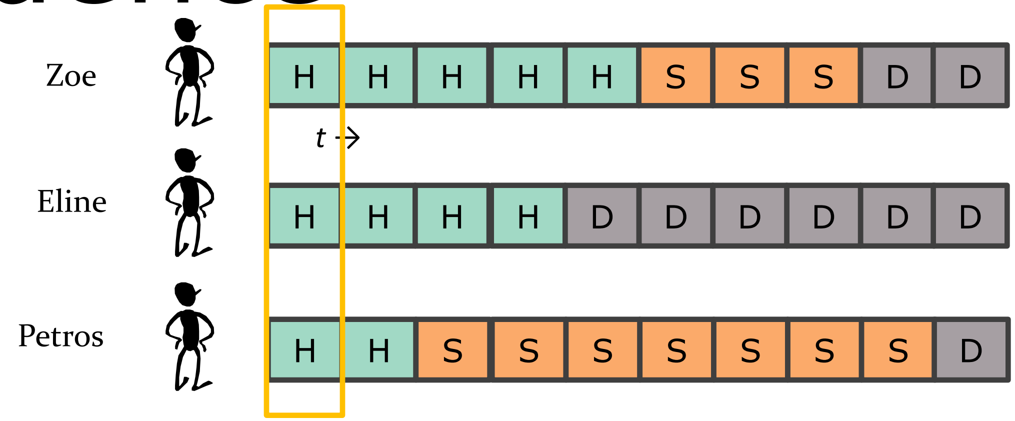
- The **probability to die** depends on the duration of being sick



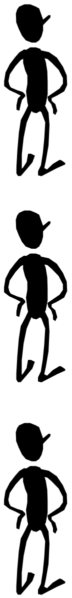
Duration of being sick (step)	p_SD
1	0.1
2	0.2
3	0.3
4	0.4
5	0.5
6+	0.7

State-residence

t = 1



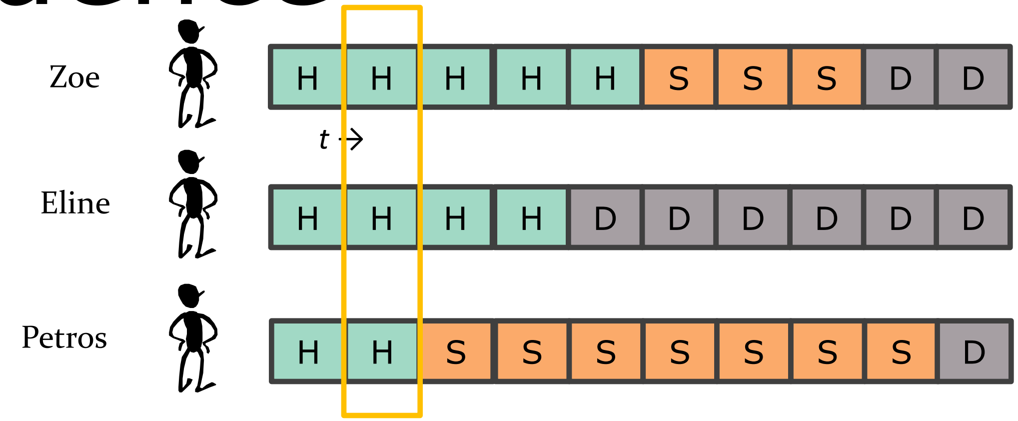
Zoe	H	H	H	H	H	S	S	S	D	D
Eline	H	H	H	H	D	D	D	D	D	D
Petros	H	H	S	S	S	S	S	S	S	D



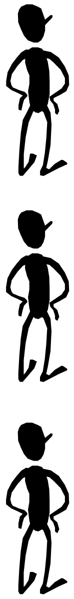
	age	sex	height	Country of birth	time Sick	p_SD
Zoe	35	Female	1.55	Taiwan	0	--
Eline	28	Female	1.68	NL	0	--
Petros	36	Male	1.89	Greece	0	--

State-residence

t = 2



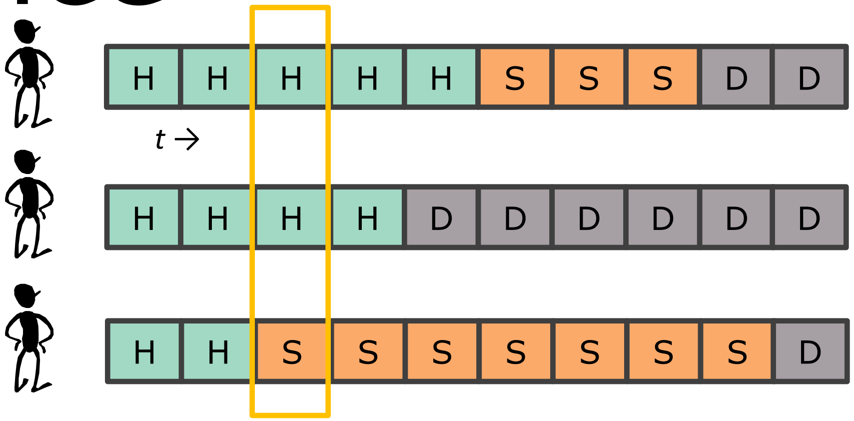
Zoe	H	H	H	H	H	S	S	S	D	D
Eline	H	H	H	H	D	D	D	D	D	D
Petros	H	H	S	S	S	S	S	S	S	D



	age	sex	height	Country of birth	time Sick	p_SD
Zoe	36	Female	1.55	Taiwan	0	--
Eline	29	Female	1.68	NL	0	--
Petros	37	Male	1.89	Greece	0	--

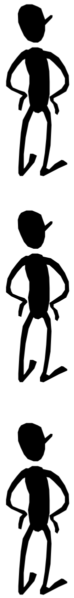
State-residence

t = 3



t →

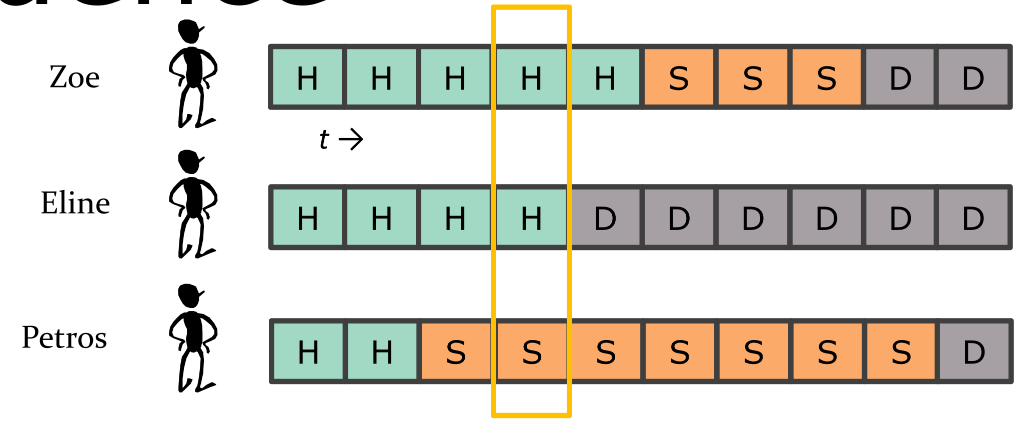
Zoe	H	H	H	H	H	S	S	S	D	D
Eline	H	H	H	H	D	D	D	D	D	D
Petros	H	H	S	S	S	S	S	S	S	D



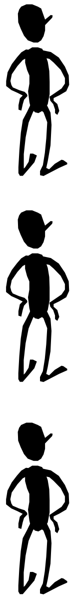
	age	sex	height	Country of birth	time Sick	p_SD
Zoe	37	Female	1.55	Taiwan	0	--
Eline	30	Female	1.68	NL	0	--
Petros	38	Male	1.89	Greece	1	0.1

State-residence

t = 4



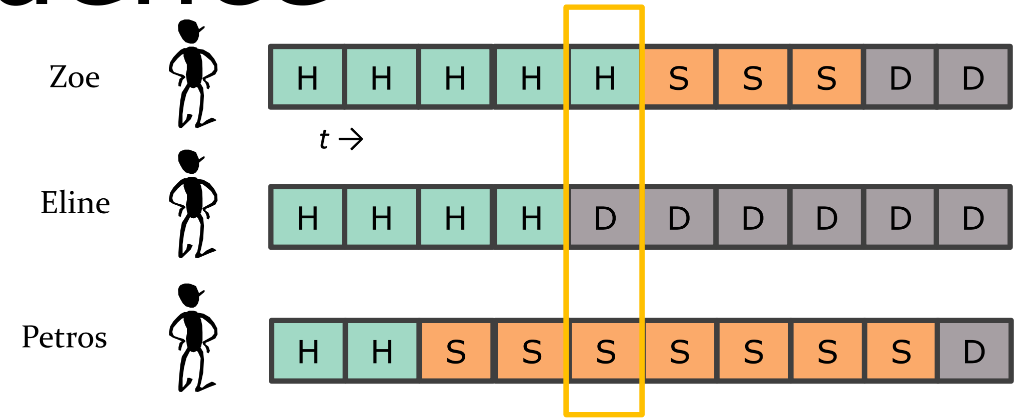
Zoe	H	H	H	H	H	S	S	S	D	D
Eline	H	H	H	H	D	D	D	D	D	D
Petros	H	H	S	S	S	S	S	S	S	D



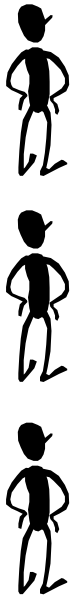
	age	sex	height	Country of birth	time Sick	p_SD
Zoe	38	Female	1.55	Taiwan	0	--
Eline	31	Female	1.68	NL	0	--
Petros	39	Male	1.89	Greece	2	0.2

State-residence

t = 5



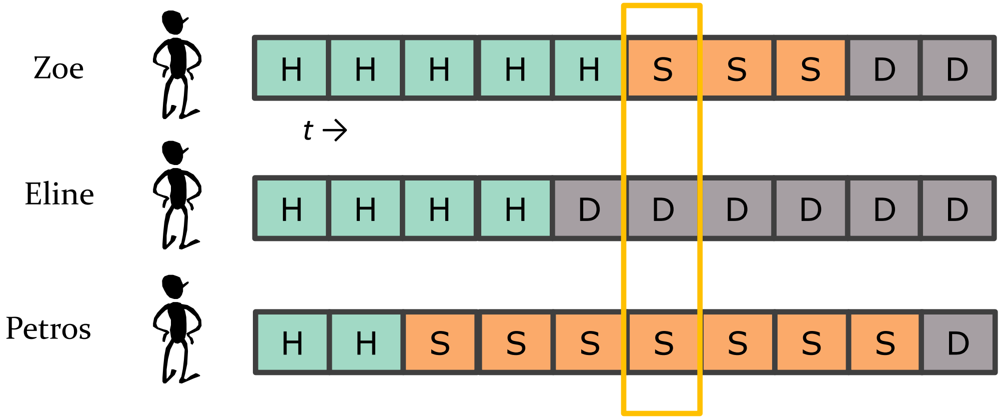
Zoe	H	H	H	H	H	S	S	S	D	D
Eline	H	H	H	H	D	D	D	D	D	D
Petros	H	H	S	S	S	S	S	S	S	D



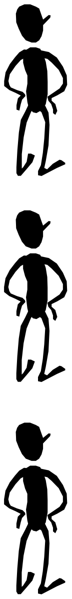
	age	sex	height	Country of birth	time Sick	p_SD
Zoe	39	Female	1.55	Taiwan	0	--
Eline	32	Female	1.68	NL	0	--
Petros	40	Male	1.89	Greece	3	0.3

State-residence

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

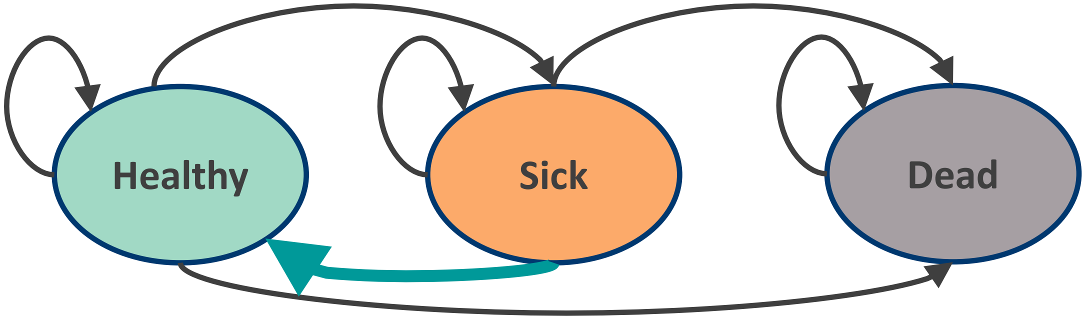


	age	sex	height	Country of birth	time Sick	p_SD
Zoe	40	Female	1.55	Taiwan	1	0.1
Eline	33	Female	1.68	NL	0	--
Petros	41	Male	1.89	Greece	4	0.4

Example (2)

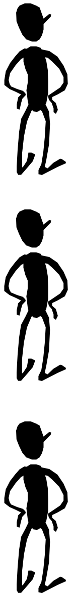
The **probability to die** still depends on the duration of being sick

3 state model where **recovery is possible**



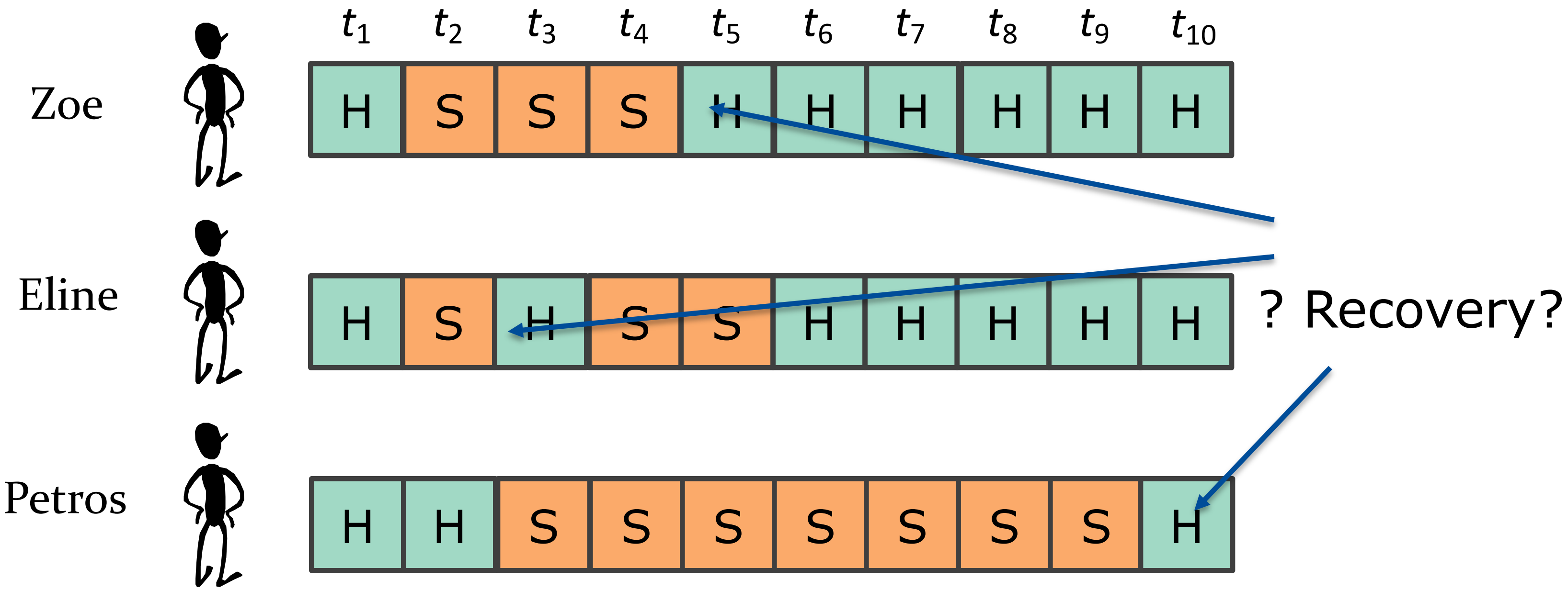
Duration of being sick (step)	**p_SD**
1	0.1
2	0.2
3	0.3
4	0.4
5	0.5
6+	0.7

New Trajectories



	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	S	H	S	S	H	H	H	H	H
Petros	H	H	S	S	S	S	S	S	S	H

Track duration of Sick



	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	S	H	S	S	H	H	H	H	H
Petros	H	H	S	S	S	S	S	S	S	H

? Recovery?

With reset: remove history after recovery

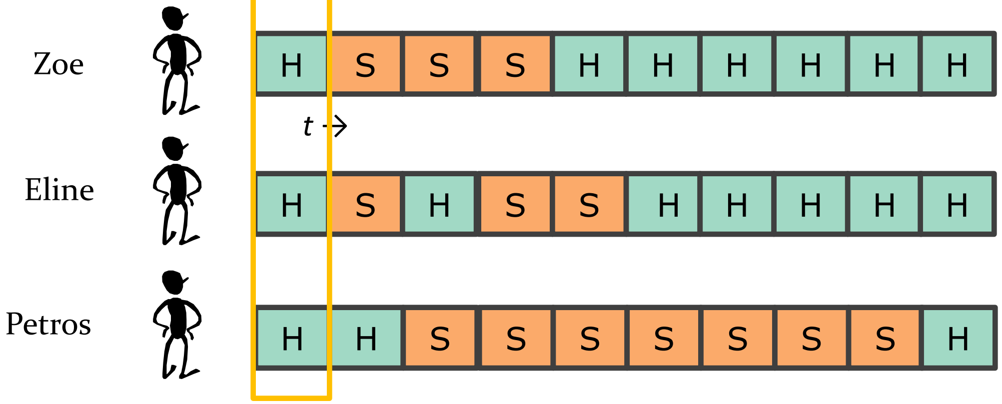
- e.g.: stop taking drugs with possible severe side effects

Without reset: history stays

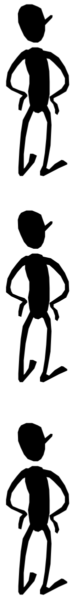
- e.g.: number of surgeries, with max

State-residence with reset

t = 1

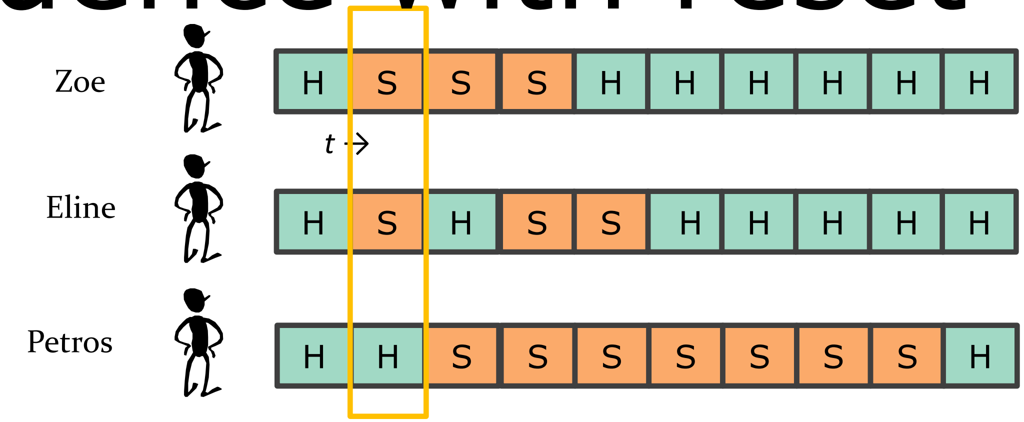


Zoe	H	S	S	S	H	H	H	H	H	H
Eline	H	S	H	S	S	H	H	H	H	H
Petros	H	H	S	S	S	S	S	S	S	H



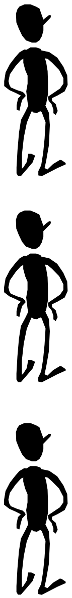
	time Sick	p_SD
Zoe:	0	--
Eline:	0	--
Petros	0	--

State-residence with reset



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

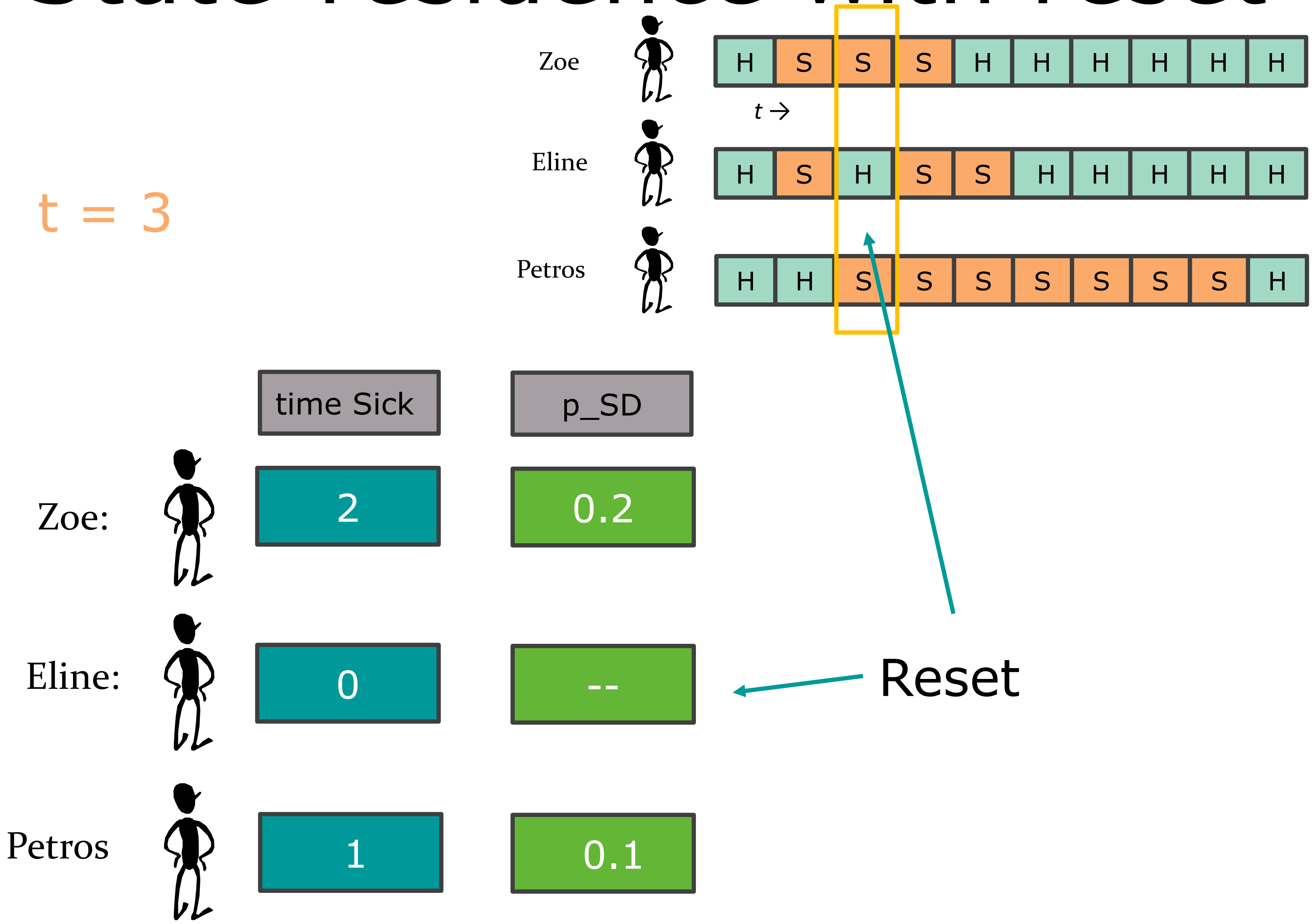
t = 2



	time Sick	p_SD
Zoe:	1	0.1
Eline:	1	0.1
Petros	0	--

State-residence with reset

t = 3



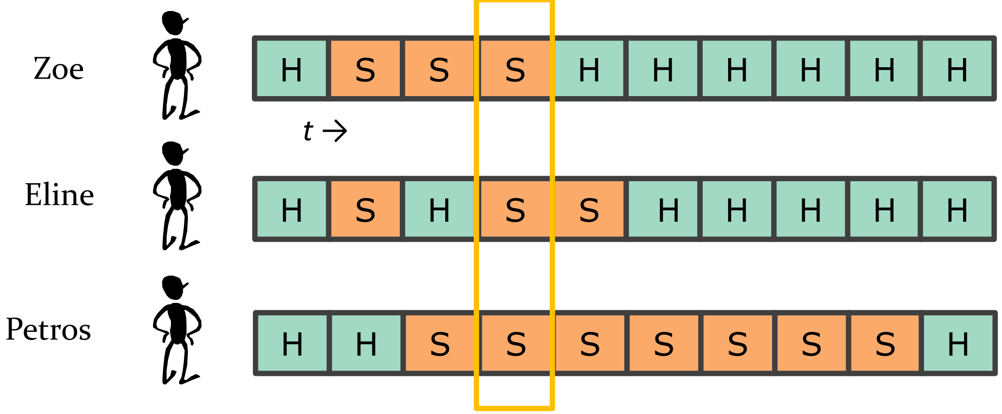
Zoe	H	S	S	S	H	H	H	H	H	H
Eline	H	S	H	S	S	H	H	H	H	H
Petros	H	H	S	S	S	S	S	S	S	H

	time Sick	p_SD
Zoe:	2	0.2
Eline:	0	--
Petros	1	0.1

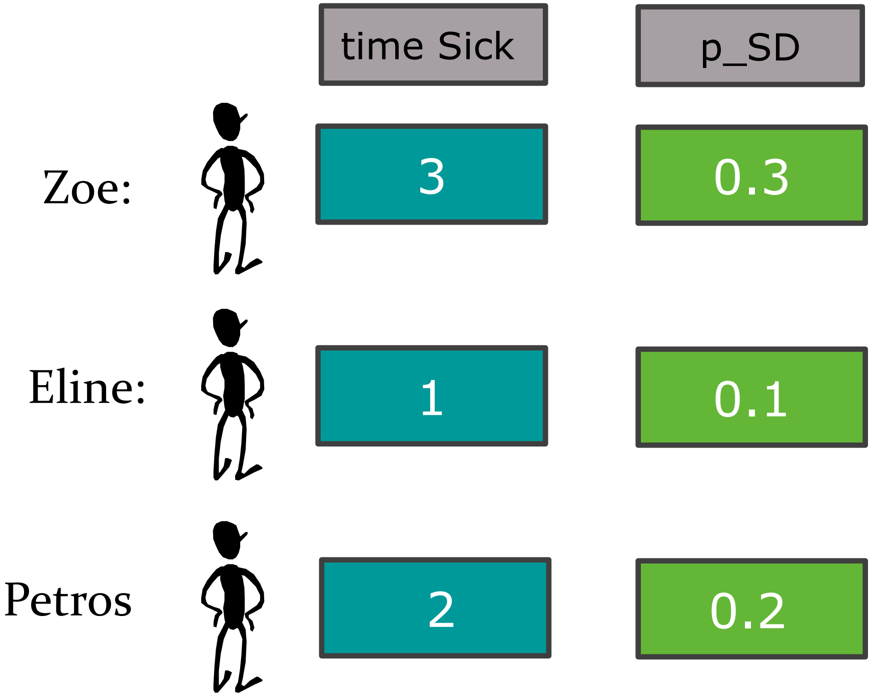
Reset

State-residence with reset

t = 4



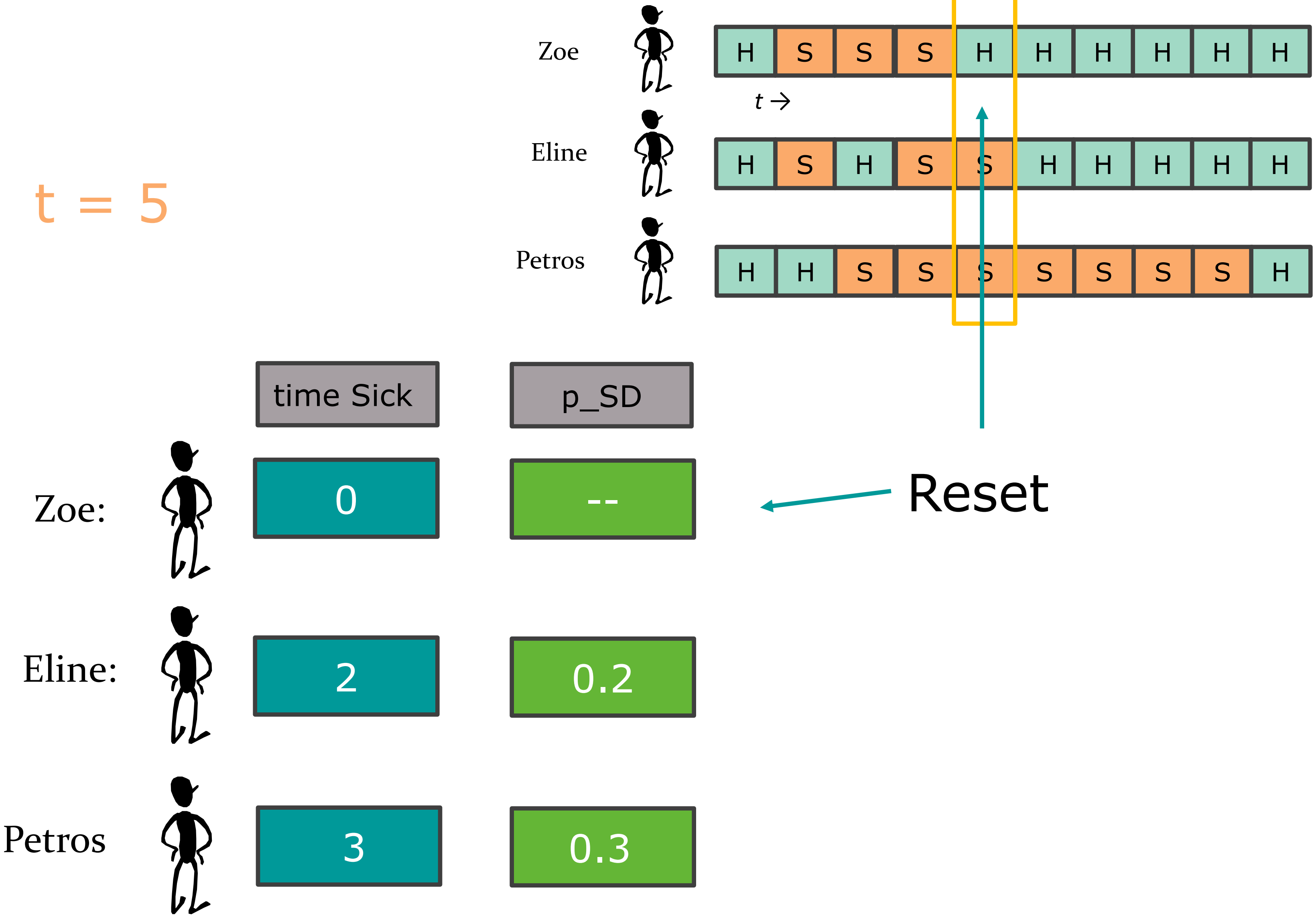
Zoe	H	S	S	S	H	H	H	H	H	H
Eline	H	S	H	S	S	H	H	H	H	H
Petros	H	H	S	S	S	S	S	S	S	H



	time Sick	p_SD
Zoe:	3	0.3
Eline:	1	0.1
Petros	2	0.2

State-residence with reset

t = 5



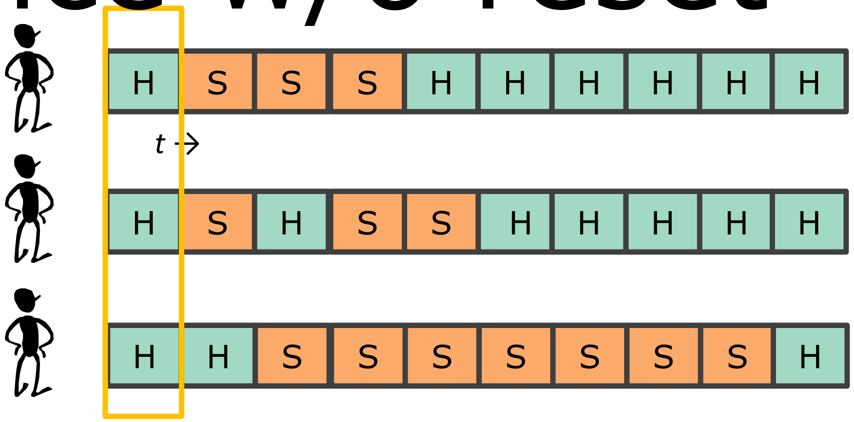
Zoe	H	S	S	S	H	H	H	H	H	H
Eline	H	S	H	S	S	H	H	H	H	H
Petros	H	H	S	S	S	S	S	S	S	H

	time Sick	p_SD
Zoe:	0	--
Eline:	2	0.2
Petros	3	0.3

Reset

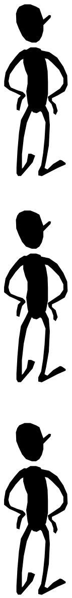
Without removing history

State-residence w/o reset



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

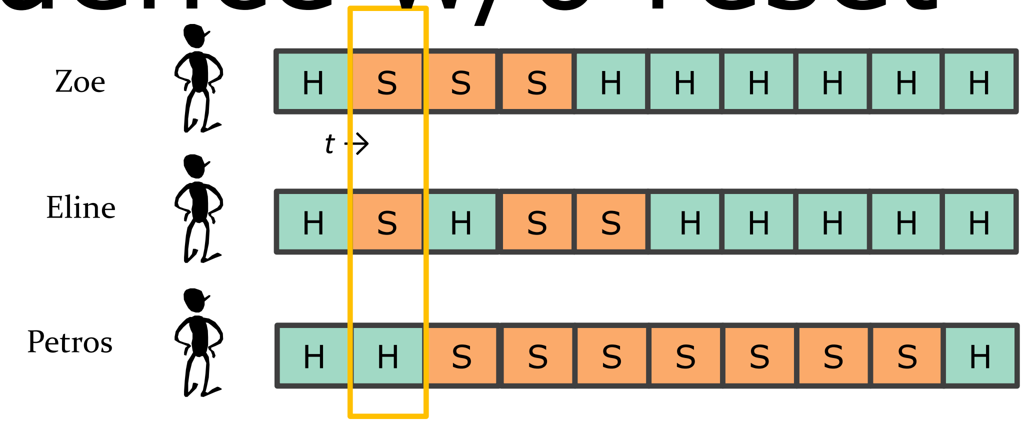
t = 1



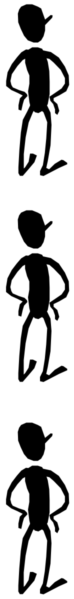
	time Sick	p_SD
Zoe:	0	--
Eline:	0	--
Petros	0	--

State-residence w/o reset

t = 2



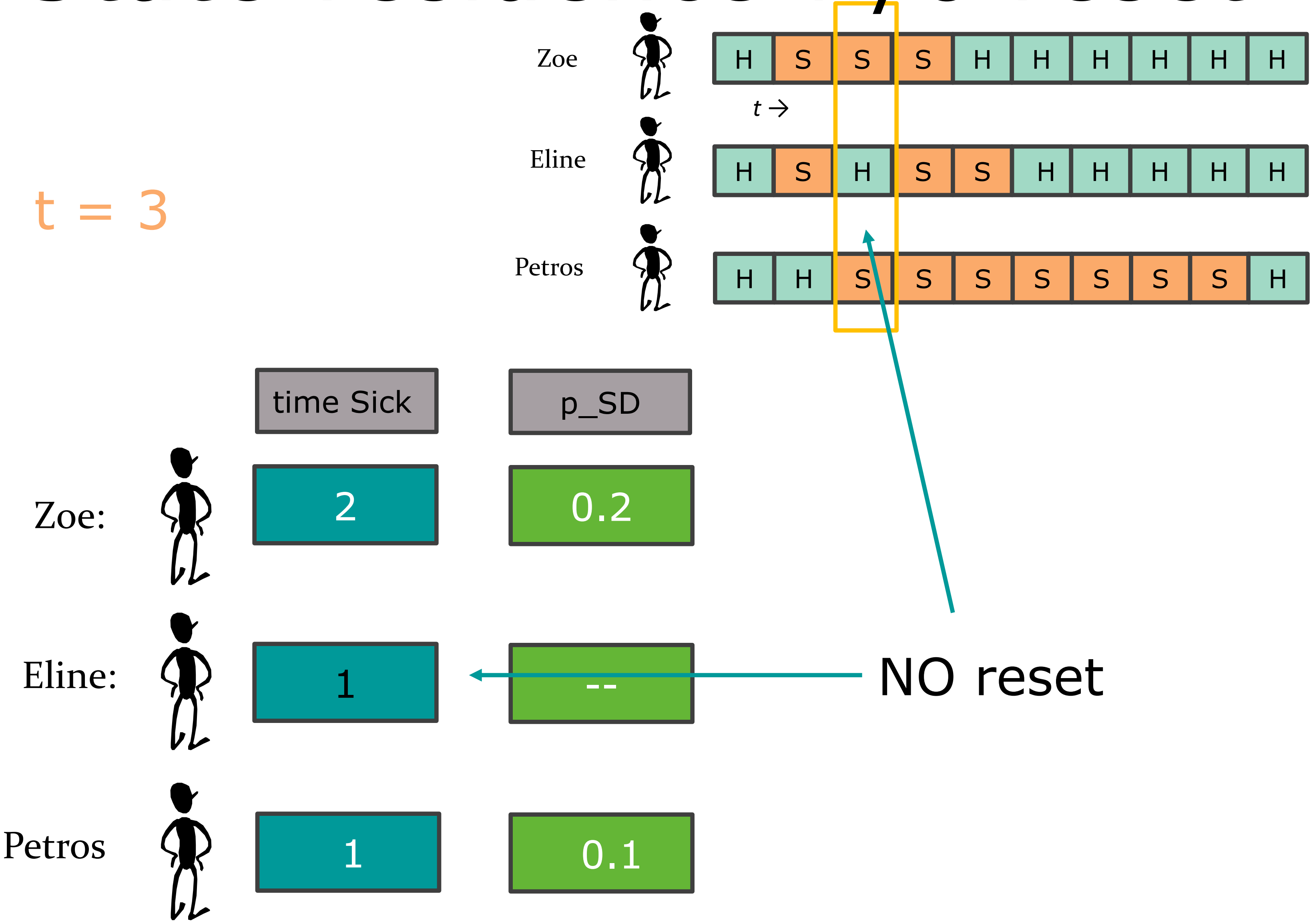
Zoe	H	S	S	S	H	H	H	H	H	H
Eline	H	S	H	S	S	H	H	H	H	H
Petros	H	H	S	S	S	S	S	S	S	H



	time Sick	p_SD
Zoe:	1	0.1
Eline:	1	0.1
Petros	0	--

State-residence w/0 reset

t = 3



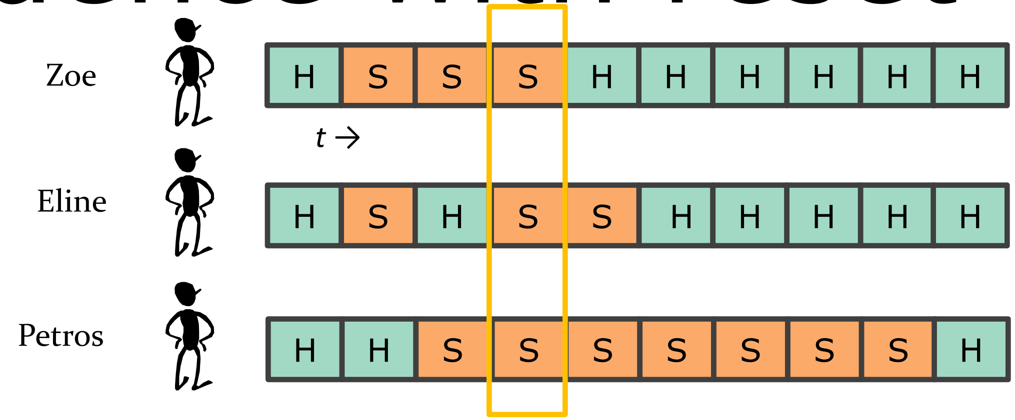
Zoe	H	S	S	S	H	H	H	H	H	H
Eline	H	S	H	S	S	H	H	H	H	H
Petros	H	H	S	S	S	S	S	S	S	H

t →

	time Sick	p_SD
Zoe:	2	0.2
Eline:	1	--
Petros	1	0.1

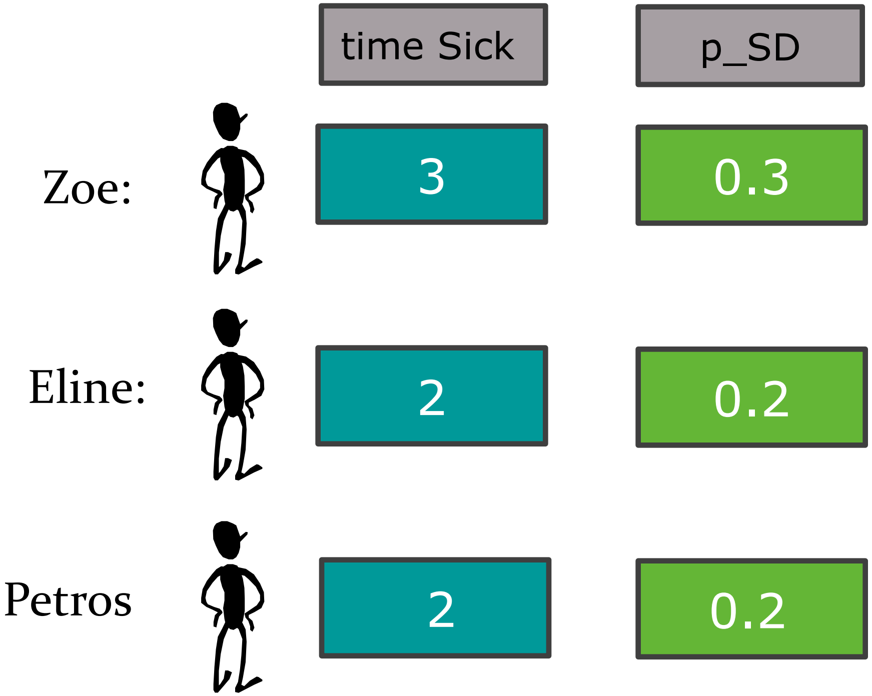
NO reset

State-residence with reset



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

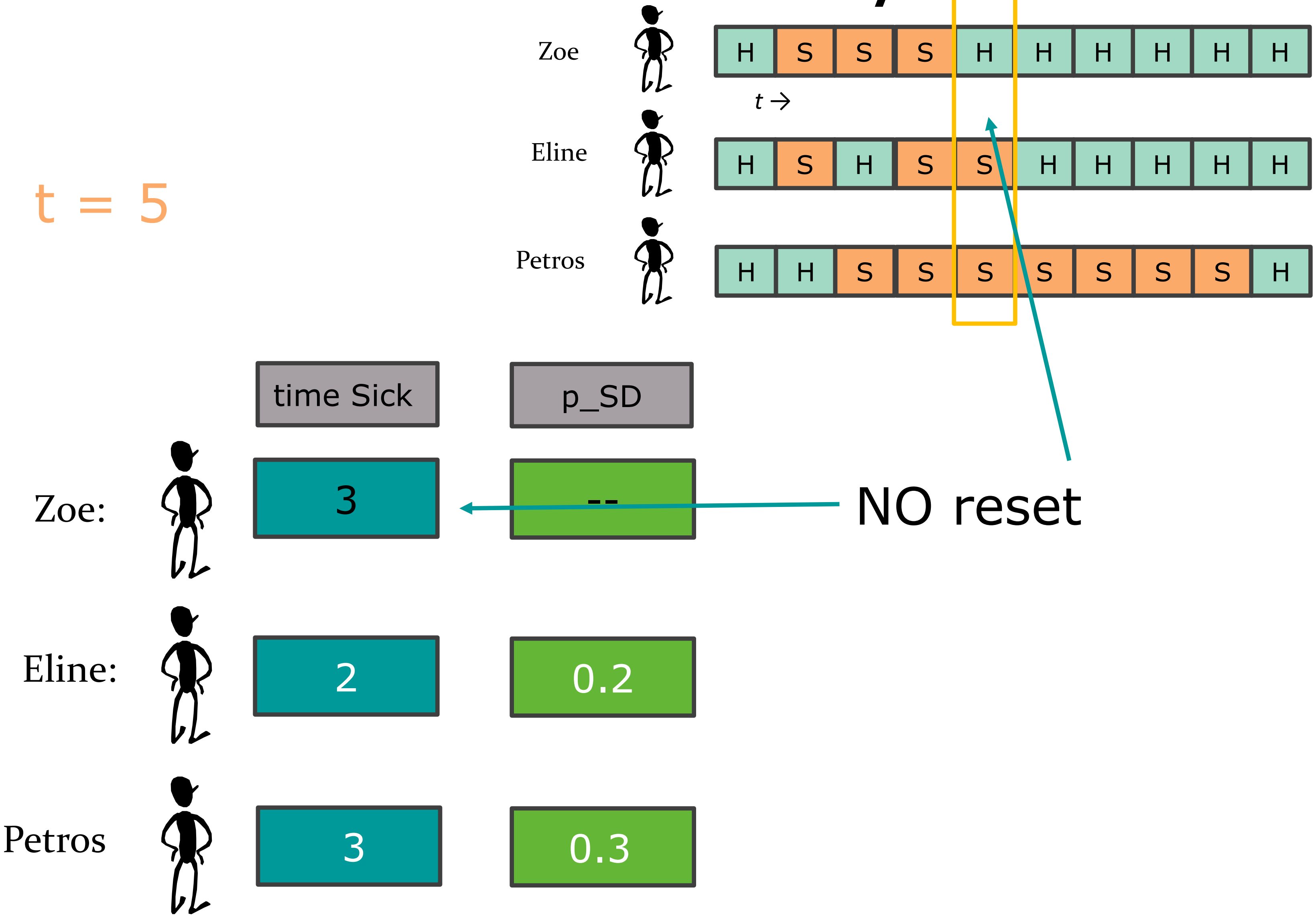
t = 4



	time Sick	p_SD
Zoe:	3	0.3
Eline:	2	0.2
Petros	2	0.2

State-residence w/o reset

t = 5



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

	time Sick	p_SD
Zoe:	3	--
Eline:	2	0.2
Petros	3	0.3

NO reset

http://darthworkgroup.com/

https://github.com/organizations/DARTH-git

https://www.linkedin.com/groups/8635339