Multi state modeling of life course trajectories

Caterina Gregorio

Trieste, 24th May 2024

1. MOX - Modelling and Scientific Computing, Department of Mathematics Politecnico di Milano, Milan, Italy
2. Biostatistics Unit, Department of Medical Sciences, University of Trieste, Trieste, Italy
3. Aging Research Center, Department of Neurobiology, Care Sciences and Society, Karolinska Institutet, Stockholm, Sweden

Outline

1. Motivation
    - Longitudinal data & multiple-events
    - Objectives and Applications in chronic illnesses
2. Competing-risk analysis
    - Definition
    - Estimation

3. Multi-state models
    - Definitions
    - Formulation
    - Time scales
    - Estimation
    - Microsimulation

4. Case-studies
    - Optimal timing for a transplant
    - Cost-effectiveness analysis
5. Conclusions

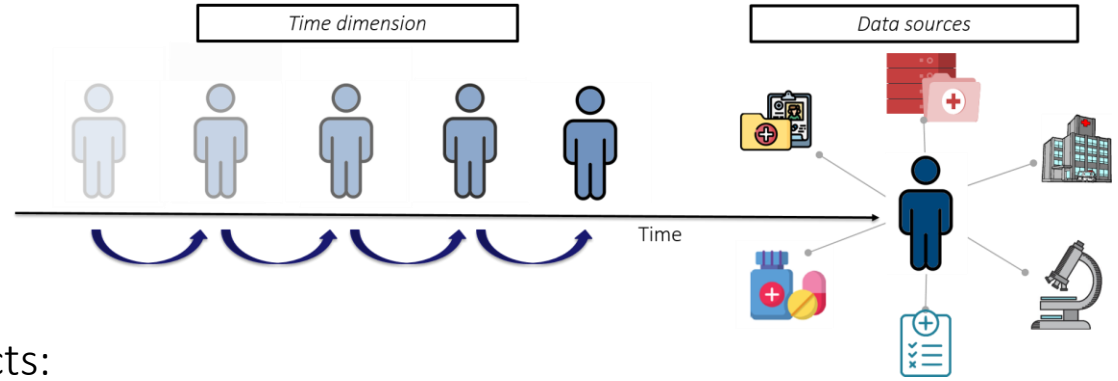
Motivation

Life courses of individuals can be defined by **multiple events** that happen to them over time. Such events determine different life trajectories depending on the **type of events** and their **timing**.

Longitudinal information
↓
Decisions

*Depending on the **phenomenon under study**, and the **population of interest**, different events can be recorded together with additional subjects' information recorded at specific times.*

Multi state models extend classical survival analysis, allowing to describe from a mathematical point of view **multiple events over time.** They represent **natural statistical methods** to model subjects' **life trajectories**.



In clinical and epidemiological research of chronic illnesses, subjects:

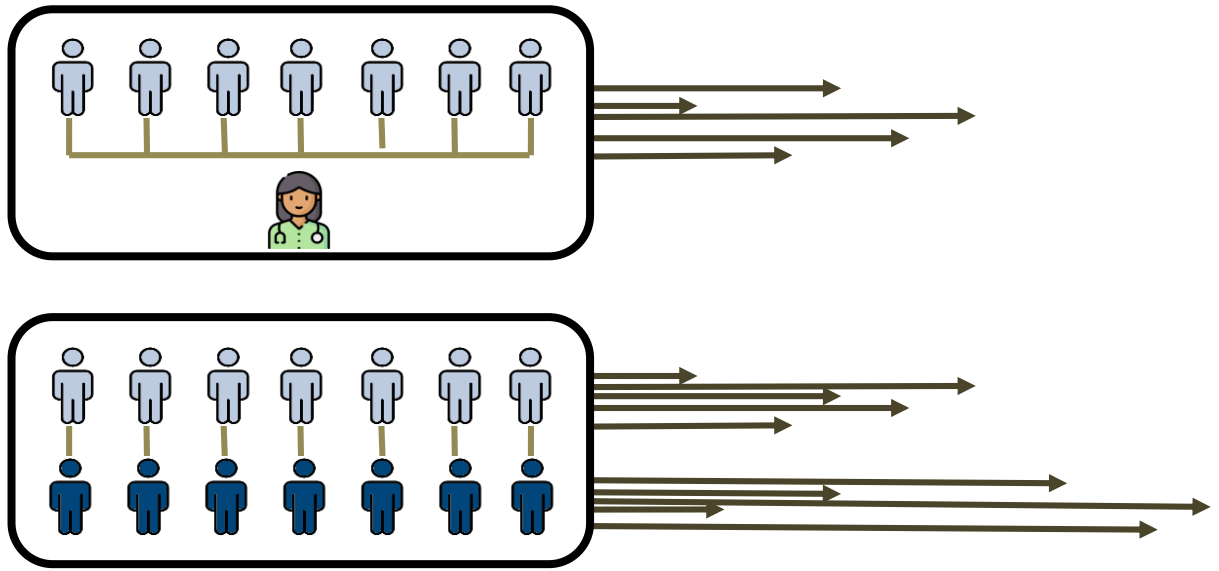
- are at risk of different type of events
- can experience multiple events of the same time
- can go through different «phases» in life

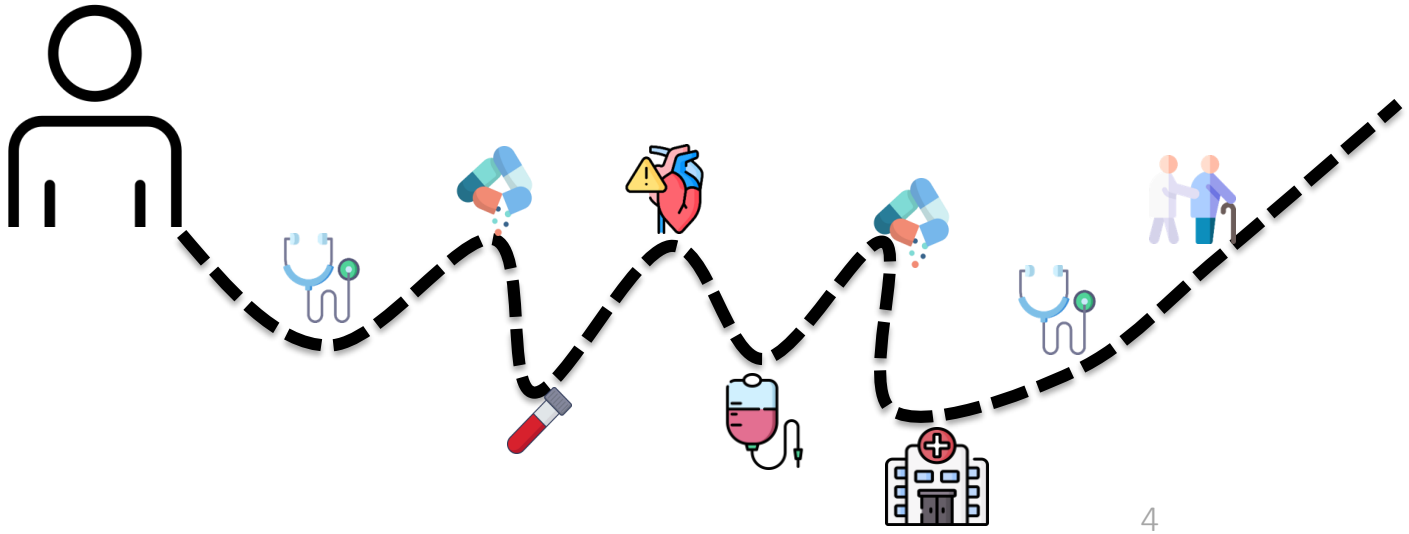
In many cases, focusing only on a single outcome is not enough …

Motivation

Some objectives of **time-to-events analyses**:

- Estimation the **effect** of a treatment/exposure
- **Cost-effectiveness** of a treatment
- **Risk Stratification** and **prediction**



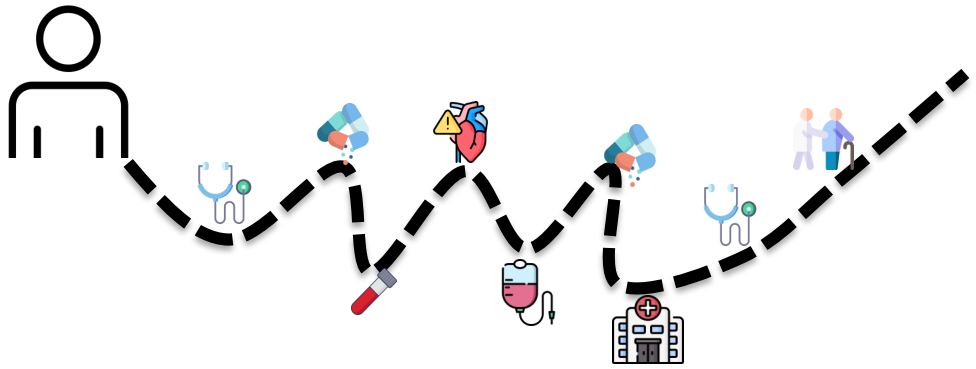


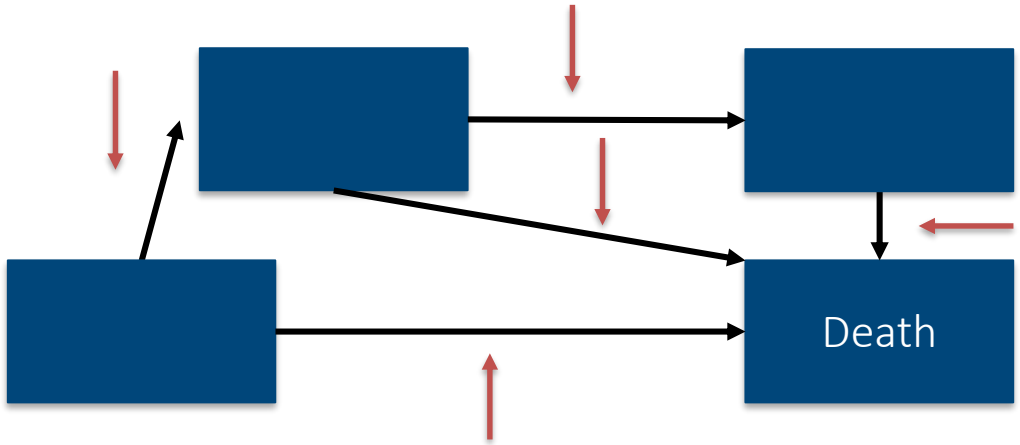
In many cases, focusing only on a single outcome is not enough …

Motivation

Some objectives of **time-to-events analyses**:

- Estimation the effect of a treatment/exposure
- Cost-effectiveness of a treatment
- Risk Stratification and prediction





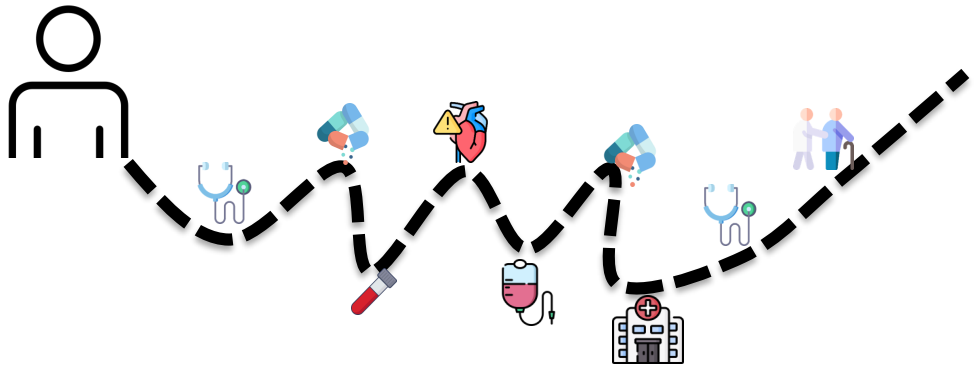
- *Exposure may impact differently each transition*
- *Long-term effect on survival needs to take into account of intermediate events*

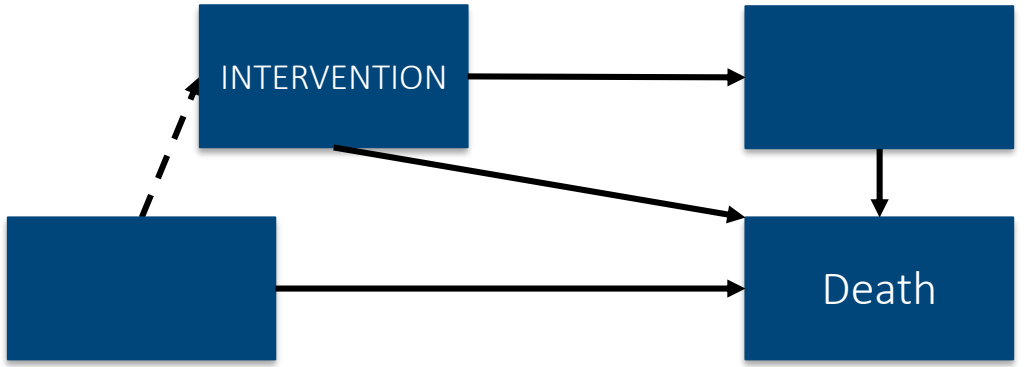
Extension to
Multi State Models

Motivation

Some objectives of **time-to-events analyses**:

- Estimation the effect of a treatment/exposure
- Cost-effectiveness of a treatment
- Risk Stratification and prediction





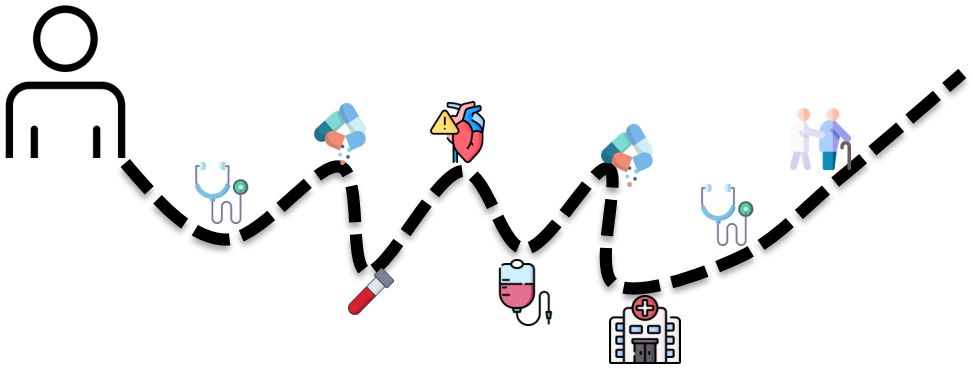
- *Exposure/treatment can itself be something that can be assigned over time*
- *It can be of interest to study the optimal timing for an intervention*

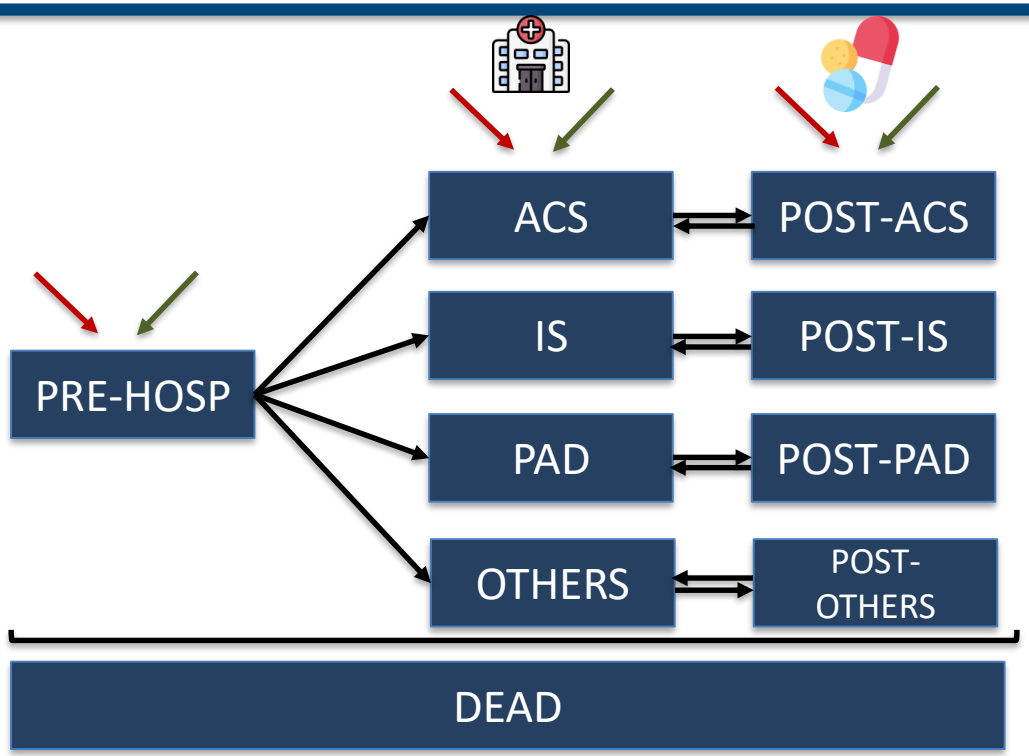
Extension to Multi State Models

Motivation

Some objectives of **time-to-events analyses**:

- Estimation the effect of a treatment/exposure
- <u>Cost-effectiveness of a treatment</u>
- Risk Stratification and prediction





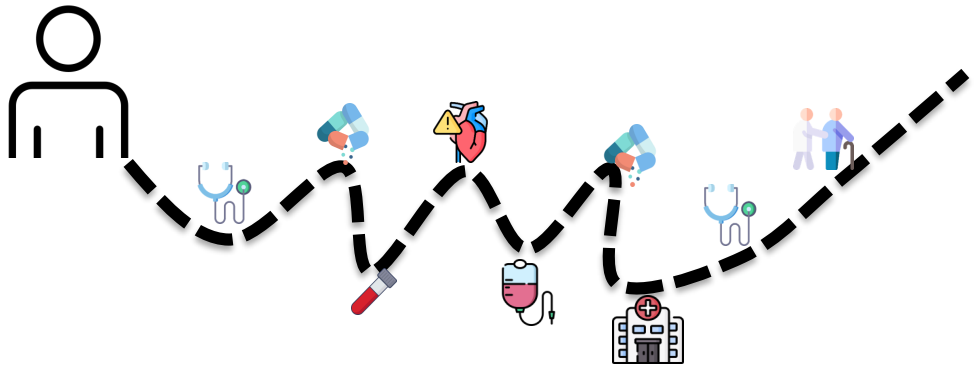
- *Cost-effectiveness needs to take into account the costs and utilities associated with each "life-phase"*

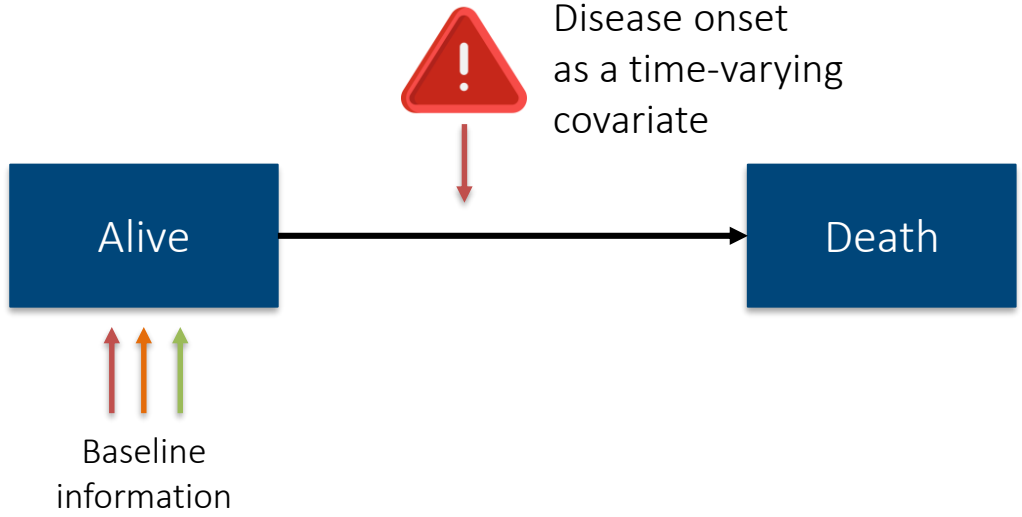
Extension to
Multi State Models

Motivation

Some objectives of **time-to-events analyses**:

- Estimation the effect of a treatment/exposure
- Cost-effectiveness of a treatment
- <u>Risk Stratification and prediction</u>





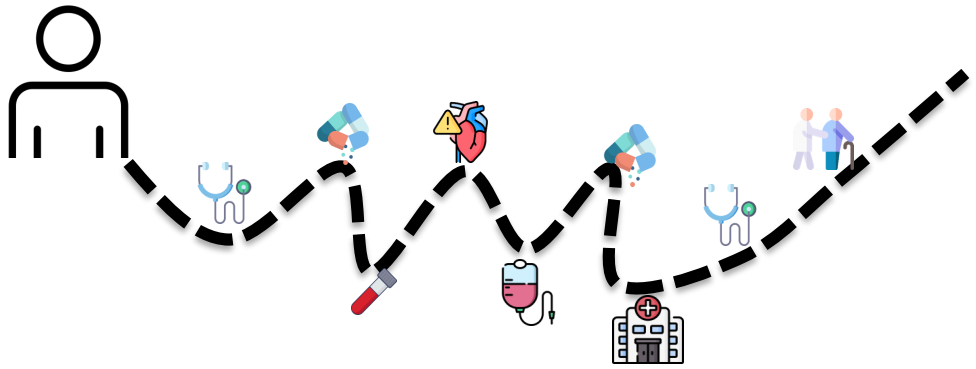
- *When predicting survival, it is also important to model the process which underlies the onset of the disease*

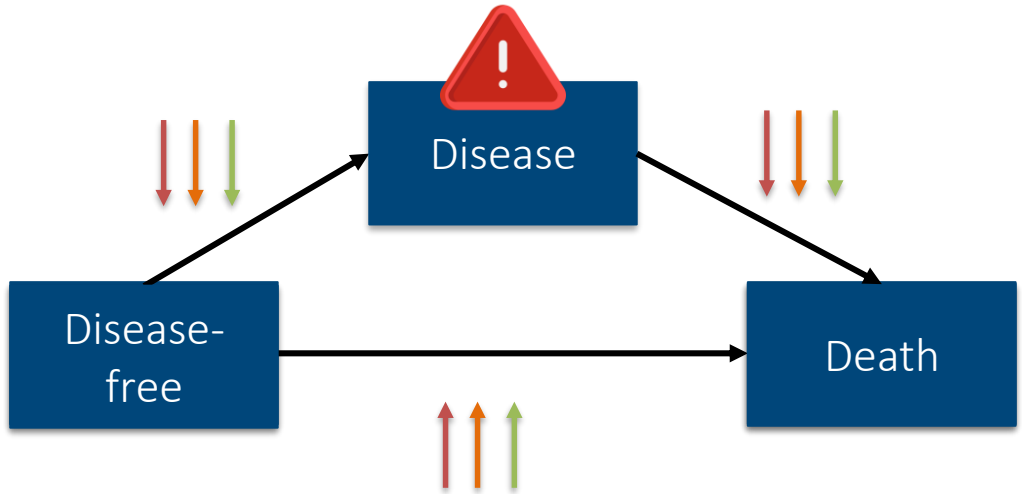
Extension to Multi State Models

Motivation

Some objectives of **time-to-events analyses**:

- Estimation the effect of a treatment/exposure
- Cost-effectiveness of a treatment
- <u>Risk Stratification and prediction</u>





- *When predicting survival, it is also important to model the process which underlies the onset of the disease*
- *The 3-states models allows to consider the process which underlies the onset of the disease*

Extension to Multi State Models

Motivation: Examples in chronic illness

- Heart Failure
- Hyperlipidemia

- Myelodysplastic syndromes

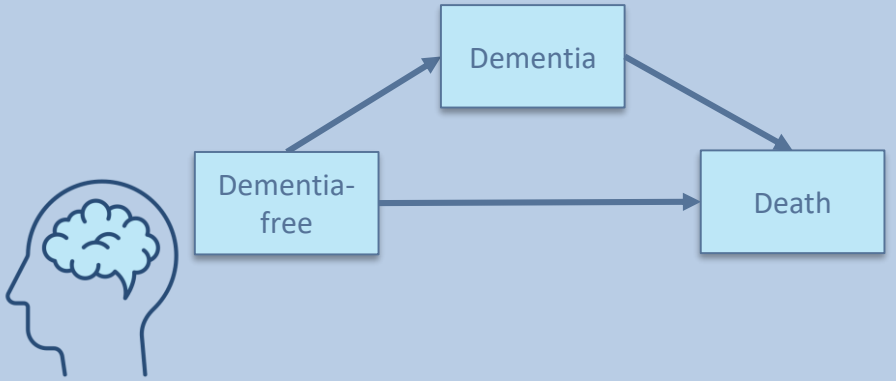
- Dementia

Motivation: Examples in chronic illness

- Heart Failure
- Hyperlipidemia

- Myelodysplastic syndromes

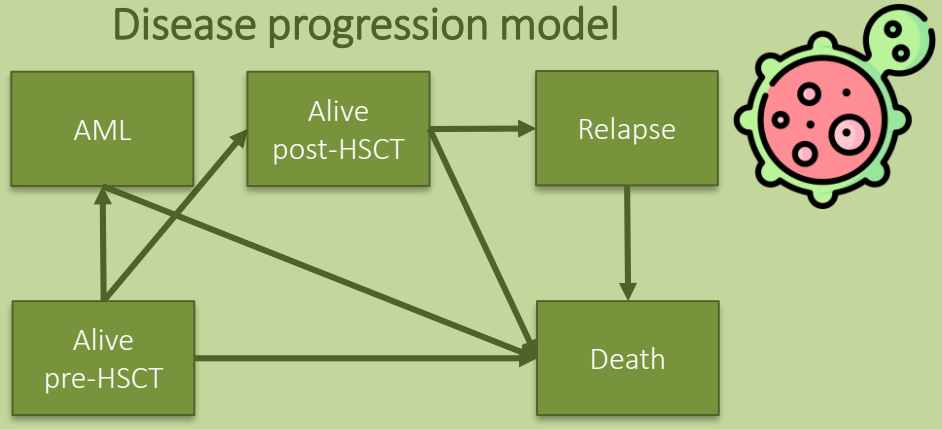
Illness-death model



Motivation: Examples in chronic illness

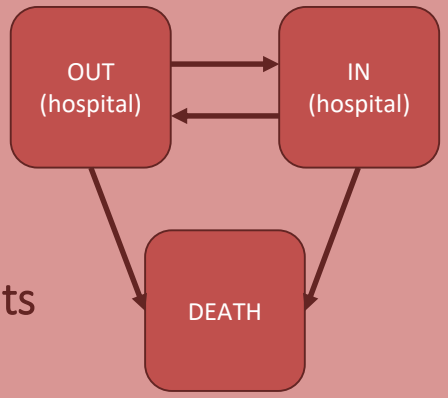
- Heart Failure
- Cardiovascular disease

Disease progression model



- Cognitive decline and dementia

Motivation: Examples in chronic illness

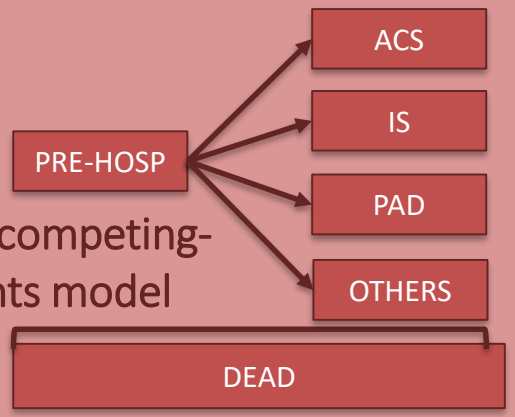


Recurrent events model

- Myelodysplastic syndromes

- Cognitive decline and dementia

Motivation: Examples in chronic illness

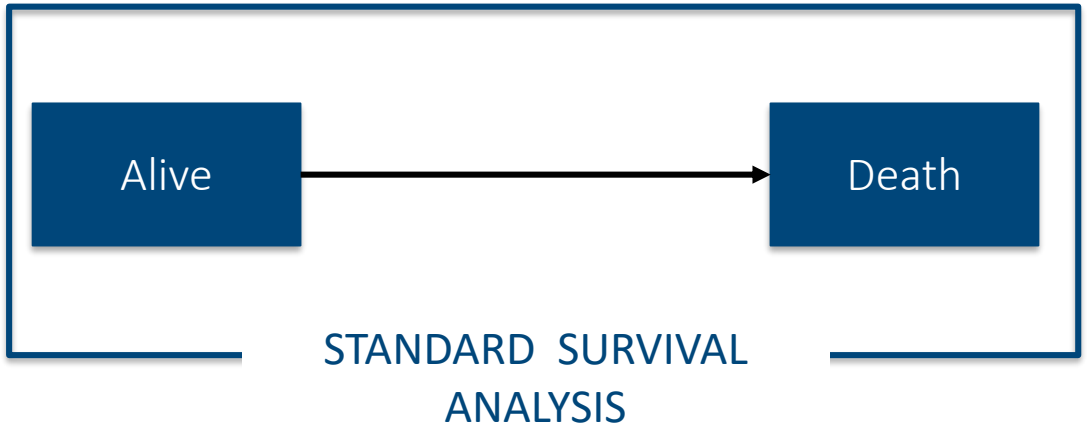


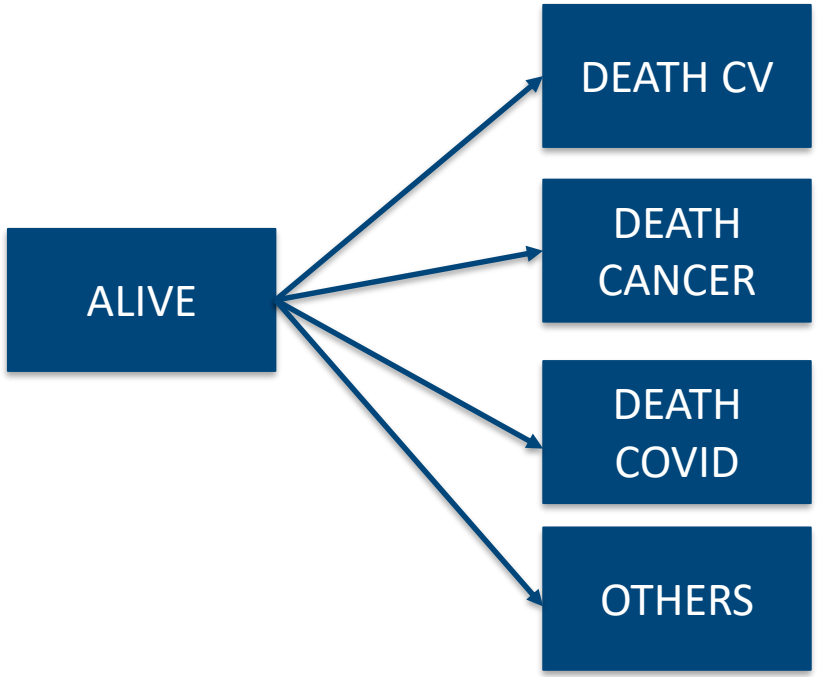
Recurrent competing-risk events model

- Myelodysplastic syndromes

- Cognitive decline and dementia

Competing risks





By definition, a **competing event** is an event that prevents one or more events from happening. Possible choices to deal with it:

- Marginal analysis → Outside of the scope of this presentation!
- Competing risk analysis

Competing risks

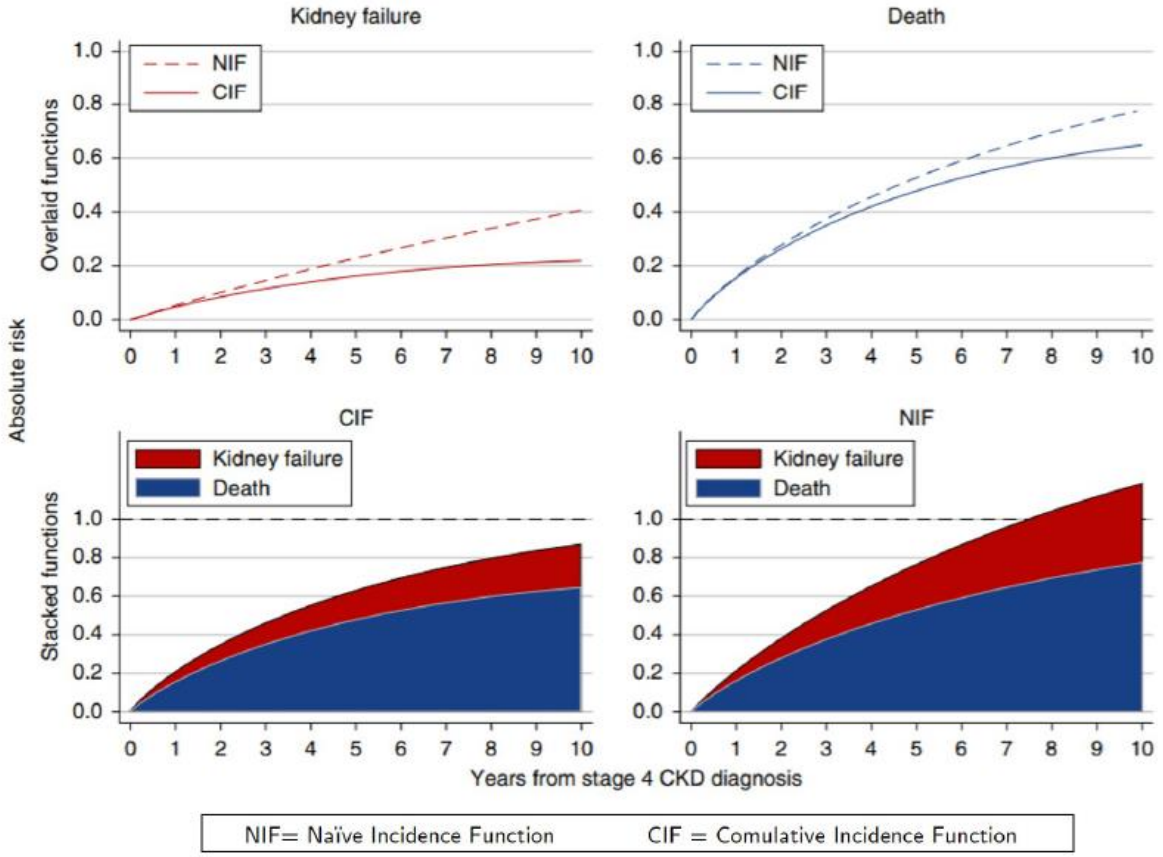
Whenever we have a setting where there are **competing events**, we need to modify the definitions of the quantities of interest for our time-to-event analysis. For each event $k = 1, \dots, K$ we now define:

CUMULATIVE INCIDENCE $$\Pr(T \leq t, E = k) = \int_0^t S(h)\lambda_k(h)dh$$

where T is the event time, E is the event type.

It depends on two other quantities:

- **Overall survival:** $S(t) = \prod_{e=1}^{K} \exp\{-\int_0^t \lambda_e(h)dh\}$
- **Cause-specific hazard:** $\lambda_k = \lim_{\Delta t \downarrow 0} \frac{\Pr(T \leq t < t+\Delta t, E = k | T \geq t)}{\Delta t}$



Competing risks

Whenever we have a setting where there are competing events, we need to modify the definitions of the quantities of interest for our time-to-event analysis. For each event $k = 1, \dots, K$ we now define:

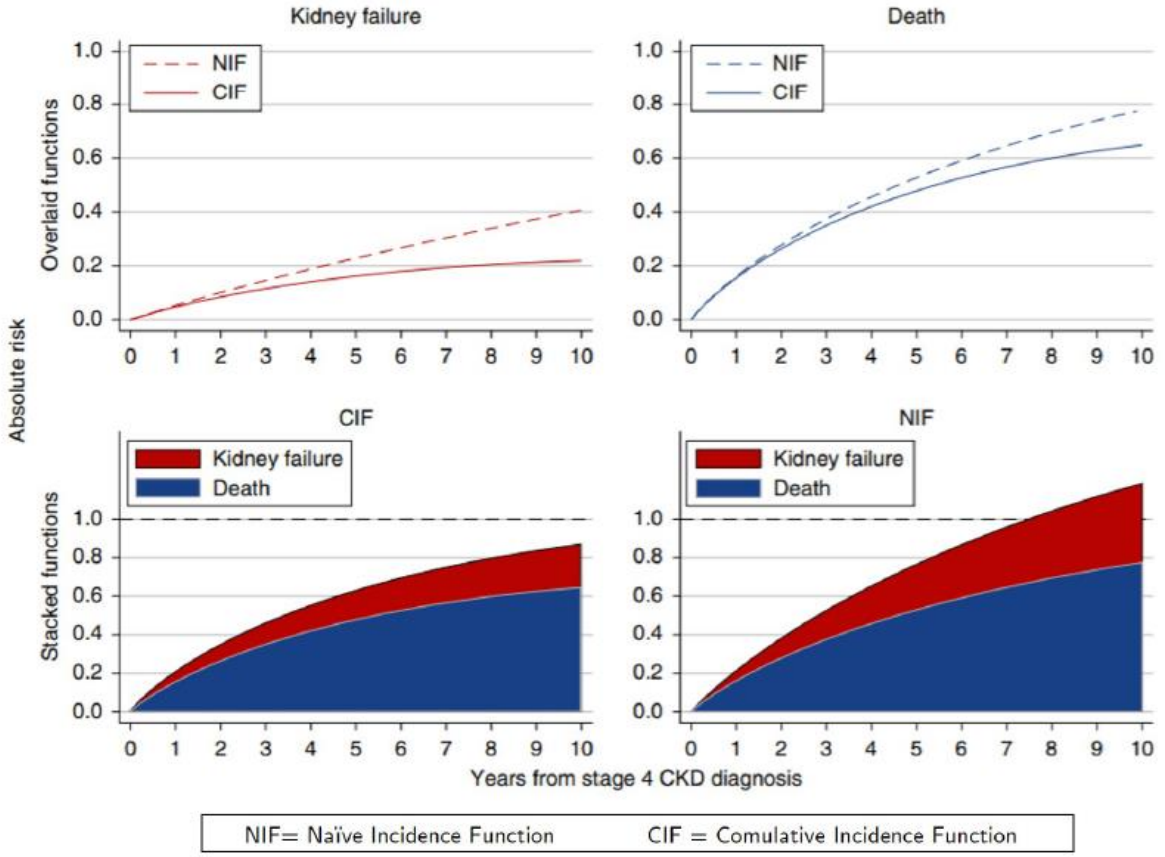
$$\Pr(T \leq t, E = k) = \int_0^t S(h)\lambda_k(h)dh$$

This quantity is equivalent to the 1-survival probability ir the standard survival setting.

It can be estimated as:

$$\int_0^t \hat{S}(h)\hat{\lambda}_k(h)dh$$

Only if K=1, we retrieve the Kaplan-Meier estimator as a special case.



Definitions

A multi state model can be described by a set of **discrete states, *S={1,…,s},*** which individuals can visit over time and by corresponding **possible transitions** between them.

Formally, an individual in the model can be denoted as a random process, $\{X(t); t > 0\}$, with $X(t)$ taking the values between ***S={1,…,s},***

Each possible transition between a state *g* and a state *h* can be defined through the **intensity function, $\lambda_{gh}(\mathrm{t})$,** which corresponds to **instantaneous probability** at each time *t* of moving from one state to the other, taking into account the **past state occupation history** up to time $t, \mathcal{H}_t = \{X(u); 0 \leq \mathrm{u} \leq \mathrm{t}\}$:

$$\lambda_{gh}(\mathrm{t}) = \lim_{\Delta t \downarrow 0} \frac{\Pr(X(t + \Delta t) = h | X(t) = g, \mathcal{H}_t)}{\Delta t}$$

The transition hazard generalizes the cause-specific hazard to the multi-state settings.

Definitions

On the cumulative scales, other quantities of interest in a multi-state model are:

- $P_{gh}(s,t) = \Pr(X(t) = h | X(s) = g, \mathcal{H}_t)$ TRANSITION PROBABILITY
- $P(X(t) = g)$ STATE OCCUPANCY PROBABILITY

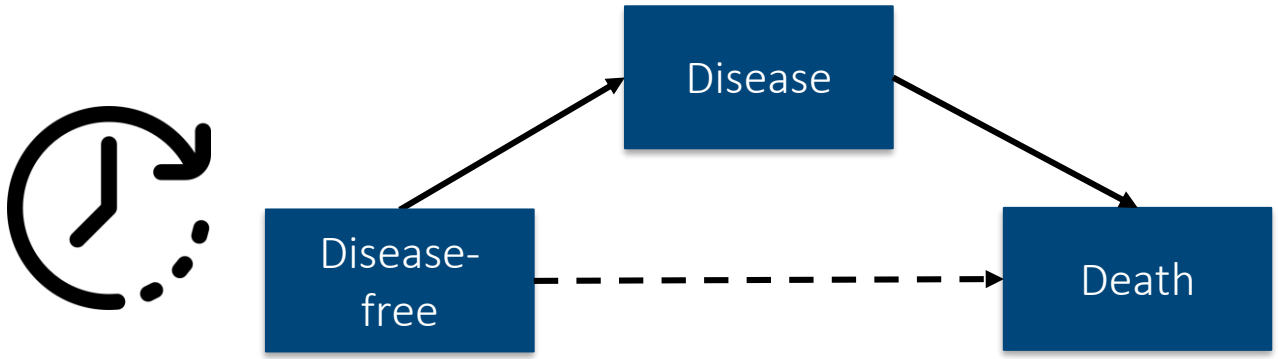
In a competing-risks model the two quantities are the same.

Others could be:

- $S(t) = 1 - Pr(X(t) = a)$ OVERALL SURVIVAL FUNCTION where a *is* the absorbing state
- $\int_o^w S(h)dh$ RESTRICTED (over horizon w) MEAN SURVIVAL TIME

Methods: time scales

A Markov multi-state model assumes that all information about the future is contained in the present state. In these models time is measured since «time 0» e.g. age at study entry. ***Clock forward***



$$P_{gh}(s,t) =$$
$$= \Pr(X(t) = h | X(s) = g, \mathcal{H}_t)$$
$$= \Pr(X(t) = h | X(s) = g)$$

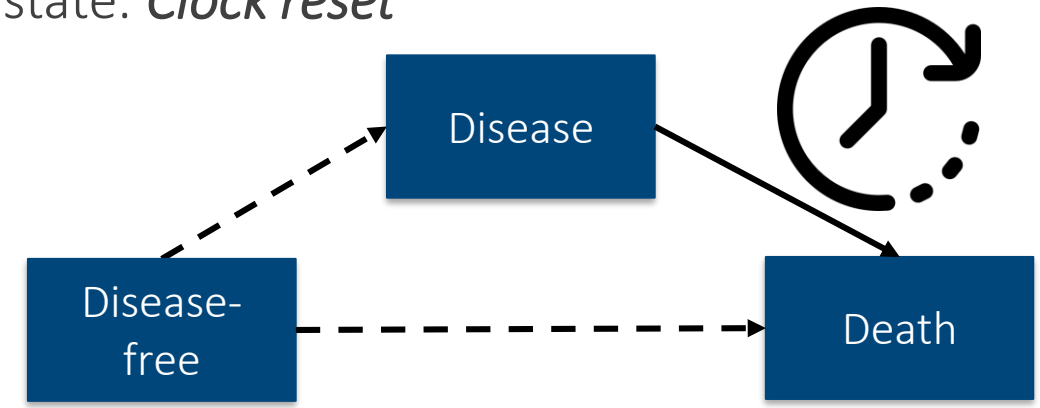
A Semi-Markov model assumes that the risk of transitioning into the next state depends on time spent in the present state. ***Clock reset***

Most of the phenomenon of interest in chronic illnesses are not markovian/semi-markovian in a strict sense but …

Methods: time scales

A Markov multi-state model assumes that all information about the future is contained in the present state. In these models time is measured since «time 0» e.g. age at study entry. ***Clock forward***

A Semi-Markov model assumes that the risk of transitioning into the next state depends on time spent in the present state. ***Clock reset***



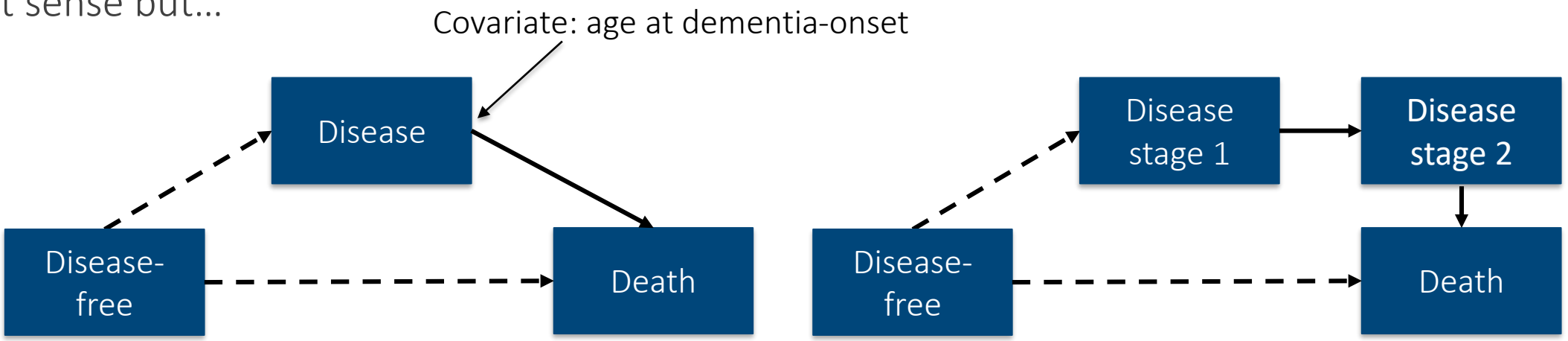
Most of the phenomeon of interest in chronic illnesses are not markovian/semi-markovian in a strict sense but …

Methods: time scales

A Markov multi-state model assumes that all information about the future is contained in the present state. In these models time is measured since «time 0» e.g. age at study entry. ***Clock forward***

A Semi-Markov model assumes that the risk of transitioning into the next state depends on time spent in the present state. ***Clock reset***

Most of the phenomenon of interest in chronic illnesses are not markovian/semi-markovian in a strict sense but…



Estimation of the hazards

We can estimate hazards from a generic state *g* to *h* using for example a generalized parametric survival model:

$$\lambda_{gh}(\mathrm{t}|\boldsymbol{x_i}) = \lambda^0_{gh}(t)\exp\{\boldsymbol{\beta^T x_i}\} \quad \textit{clock-forward approach}$$
$$\lambda_{gh}(\mathrm{u}|\boldsymbol{x_i}) = \lambda^0_{gh}(u)\exp\{\boldsymbol{\beta^T x_i}\} \quad \textit{clock-reset approach}$$

→The baseline hazard is estimated together with the coefficients
→More flexibility in the shape of the baseline hazards, only on the number of knots
→It can be easily extended to include time-varying effects for the covariates

Specifically, the logarithm of the baseline cumulative hazard function is modelled as a natural cubic spline function of log time:

$$logH_{gh}(t|\boldsymbol{x_i}) = logH^0_{gh}(t) + \boldsymbol{\beta}^T\boldsymbol{x_i} = b^m_{gh}(log(t);\boldsymbol{\gamma}) + \boldsymbol{\beta^T x_i}$$

where $b^m_{gh}(\log(t);\boldsymbol{\gamma})$ is a natural cubic spline with m knots. The parameters in the model can be estimated using Maximum Likelihood (ML) and their uncertainly can be evaluated using standard ML asymptotic theory.

Estimation of the hazards

We can estimate hazards from a generic state *g* to *h* using for example a generalized parametric survival model:

$$\lambda_{gh}(\mathrm{t}|\boldsymbol{x_i}) = \lambda^0_{gh}(t)\exp\{\boldsymbol{\beta^T x_i}\} \quad \textit{clock-forward approach}$$
$$\lambda_{gh}(\mathrm{u}|\boldsymbol{x_i}) = \lambda^0_{gh}(u)\exp\{\boldsymbol{\beta^T x_i}\} \quad \textit{clock-reset approach}$$

Regardless of whether the models are parametrically or semi-parametrically specified**, the estimation can be obtain using standard software assuming that the format of the dataset is correct**.

The model changes according to the structure of the dataset. Basic rules for constructing the dataset in the correct format are (also called **counting-process formulation**):

- Each **transition** will have its **own sub-dataset**
- Each time a subject **start being "at risk"** for a transition a new row needs to be created considering:
  - time for which subject **started being at risk** for that transition
  - time for which subject **stopped being at risk** for that transition
  - at the end of the "at-risk" interval **did it experience the transition**?
- A subject can be at risk of multiple events at the same time but only one can be experience in the same interval.
- Every time subject experience an event (transition to a different state) he enters a "new interval" so possibly multiple new rows needs to be created (unless the state entered is an absorbing state).

An Example: it easier done than said!

A B C

A, B and C enter the study when they are diagnosed with Heart Failure at age 60, 73 and 80 respectively.

It is of interest to study the **time to death** and by the administrative end of our study, 11 years later, we know that A is still alive, B and C died aged 80 and 90 respectively.

STANDARD SURVIVAL ANALYSIS

ID	Tstart	Tstop	time	status	trans
A	0	11	11	0	DEATH
B	0	7	7	1	DEATH
C	0	10	10	1	DEATH

- One possible transition
- All subjects enter at time 0 so time since entering is equal to the time spent in the state.
- Each subject has only one row in the dataset

It this the only possible scale to measure time?

An Example: it easier done than said!

A B C

A, B and C enter the study when they are diagnosed with Heart Failure at age 60, 73 and 80 respectively.

It is of interest to study the **time to death for heart failure** and by the administrative end of our study, 11 years later, we know that A is still alive, B and C died aged 80 and 90 respectively. We also know that B died of cancer while C died because of its underlying heart failure.

COMPETING RISK ANALYSIS

ID	Tstart	Tstop	time	status	trans
A	0	11	11	0	HF DEATH
B	0	7	7	0	HF DEATH
C	0	10	10	1	HF DEATH

ID	Tstart	Tstop	time	status	trans
A	0	11	11	0	O DEATH
B	0	7	7	1	O DEATH
C	0	10	10	0	O DEATH

- Two possible transitions→ two sub-datasets
- All subjects enter at time 0 so time since entering is equal to the time spent in the state.
- Each subject has one row in each sub-dataset

Models can be either estimated separately or together depending on what we think they have in common…

An Example: it easier done than said!

A B C

A, B and C enter the study when they are diagnosed with Heart Failure at age 60, 73 and 80 respectively.

It is of interest to study the **time to death taking into account the onset of chronic kidney disease (ckd)** and by the administrative end of our study, 11 years later, we know that A is still alive, B and C died aged 80 and 90 respectively. We also know that B died of cancer while C died because of its underlying heart failure. Subjects A and B were diagnosed with ckd age 70 and 79 respectively.

ILLNESS-DEATH MODEL

ID	Tstart	Tstop	time	status	Trans
A	0	10	10	1	CKD
B	0	6	6	1	CKD
C	0	10	10	0	CKD

- Three possible transitions→ three sub-datasets
- For death after ckd, we need to choose a time scale for our model.
- Each subject has either 1 or 0 row per sub-dataset.

ID	Tstart	Tstop	time	Status	trans
A	0	10	10	0	DEATH PRE
B	0	6	6	0	DEATH PRE
C	0	10	10	1	DEATH PRE

ID	Tstart	Tstop	time	Status	trans
A	10	11	1	0	DEATH POST
B	6	7	1	1	DEATH POST

An Example: it easier done than said!

A B C

A, B and C enter the study when they are diagnosed with Heart Failure at age 60, 73 and 80 respectively.

It is of interest to study the **time to death taking into hospitalizations** and by the administrative end of our study, 11 years later, we know that A is still alive, B and C died aged 80 and 90 respectively. We know that A was hospitalized at age 61.3 and it stayed in hospital 20 days while subject C was hospitalized at age 83.4 (10 days) and 89.8 (died in hospital).

REPEATED EVENTS AND DEATH

ID	Tstart	Tstop	time	status	Trans
A	0	1.3	1.3	1	IN
A	1.35	11	9.65	0	IN
B	0	7	7	0	IN
C	0	3.4	3.4	1	IN
C	3.7	9.8	6.1	1	IN

ID	Tstart	Tstop	time	status	trans
A	1.3	1.35	0.05	1	OUT
C	3.4	3.7	0.03	1	OUT
C	9.8	10	0.02	0	OUT

ID	Tstart	Tstop	time	status	Trans
A	0	1.3	1.3	0	OUT-D
A	1.35	11	9.65	0	OUT-D
B	0	7	7	1	**OUT-D**
C	0	3.4	3.4	0	**OUT-D**
C	3.7	9.8	6.1	0	**OUT-D**

ID	Tstart	Tstop	time	status	trans
A	1.3	1.35	0.05	0	IN-D
C	3.4	3.7	0.03	0	IN-D
C	9.8	10	0.02	1	IN-D

Estimation of the hazards

- Creating the dataset is intuitive, yet **time consuming** with a large sample size.
- Depending on the structure of the dataset there are **different functions in R** which can help, even though for general structures/repeated events hard coding is sometimes still necessary.

- As the number of transitions increase, of course more **choices regarding the modelling needs to be made**:
  - Time-scale
  - Baseline hazards
  - "Similarities" between transitions

- Even the **structure of the multi state model** can depend on assumptions and modelling choices

  E.g. the first hospitalization is the same as the second one?

- How to decide when my estimated model is good enough? Standard methods are valid:
  - ✓ AIC and BIC
  - ✓ Comparison with non-parametric estimates
  - ✓ Residuals

GOODNESS OF FIT

Estimating… all the rest

Estimation of the cumulative quantities are a matter algebra under the Markov assumption, once the transition-hazards have been estimated.

$$\hat{P}_{gh}(s,t) = \prod_{s \le u \le t} \{ \mathrm{I} + \widehat{d\Lambda(u)} \}$$ AALEN-JOHANSEN ESTIMATOR

Where $\widehat{d\Lambda(u)}$ is the KxK matrix that has transition hazards on off-diagonal elements while on the diagonal it has minus the sum of all the off-diagonal elements from the same row.

For state-occupancy probabilities, we can use the same estimator since they correspond to transition probabilities from the initial state.

The formula for the cumulative incidence function in the competing risks setting is also special case of this formula.

Microsimulation

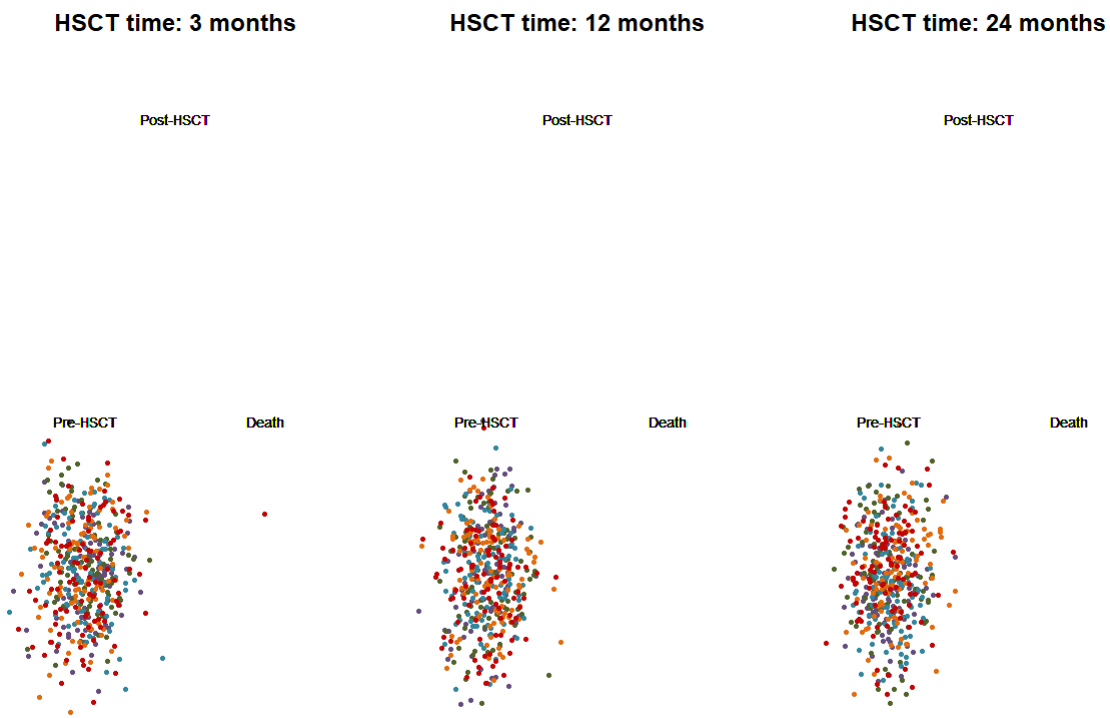
What is (dynamic) microsimulation?

They are **individual based** simulations from an **estimated multi-state model**.

They simulate **individual trajectories** in continuous time between **health states using random number generation**.

Why to simulate?

- **Cumulative quantities** can be analytically obtained only under markov assumption and the computation can be tedious as the model move away from the simple illness-death model
- Perform **scenario analyses**



Microsimulation

What is microsimulation?

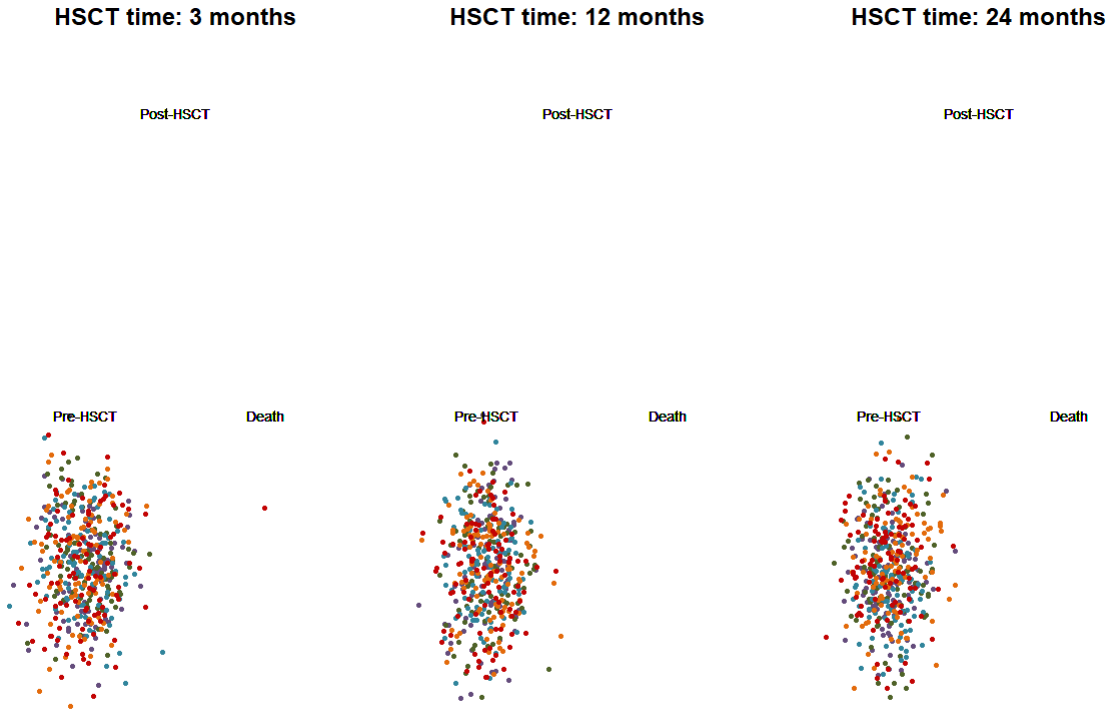
The term microsimulation was introduced by **Orcutt** in **1957** in the context of economic applications.

Since then, they have been used also extensively in **health applications**.

From a statistical perspective, they belong to the class of **Monte Carlo simulation methods**.

Why to simulate?

- **Cumulative quantities** can be analytically obtained only under markov assumption and the computation can be tedious as the model move away from the simple illness-death model
- Perform **scenario analyses**

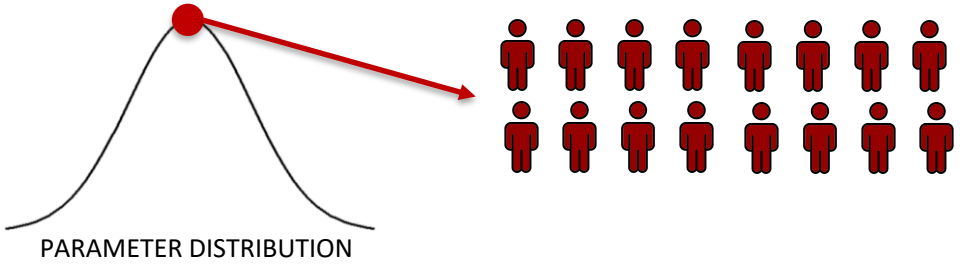


More on microsimulations: Probability Sensibility Analysis

- It is used take into account uncertainty in the estimations of the parameters of the disease model into the microsimulation
- It is inherently a Bayesian idea even tough it can be used also under a frequentist framework

STANDARD MICROSIMULATION

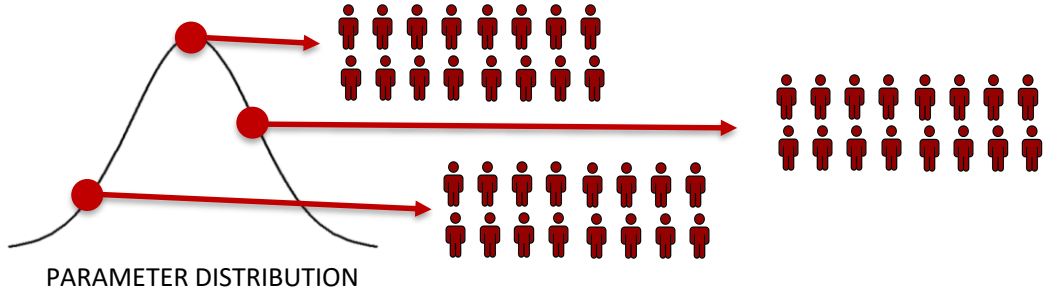
Individuals are simulated using the same value for each parameter



- Disease model→ MLE point estimates

PSA MICROSIMULATION

Two levels of simulations: Parameters → Individuals

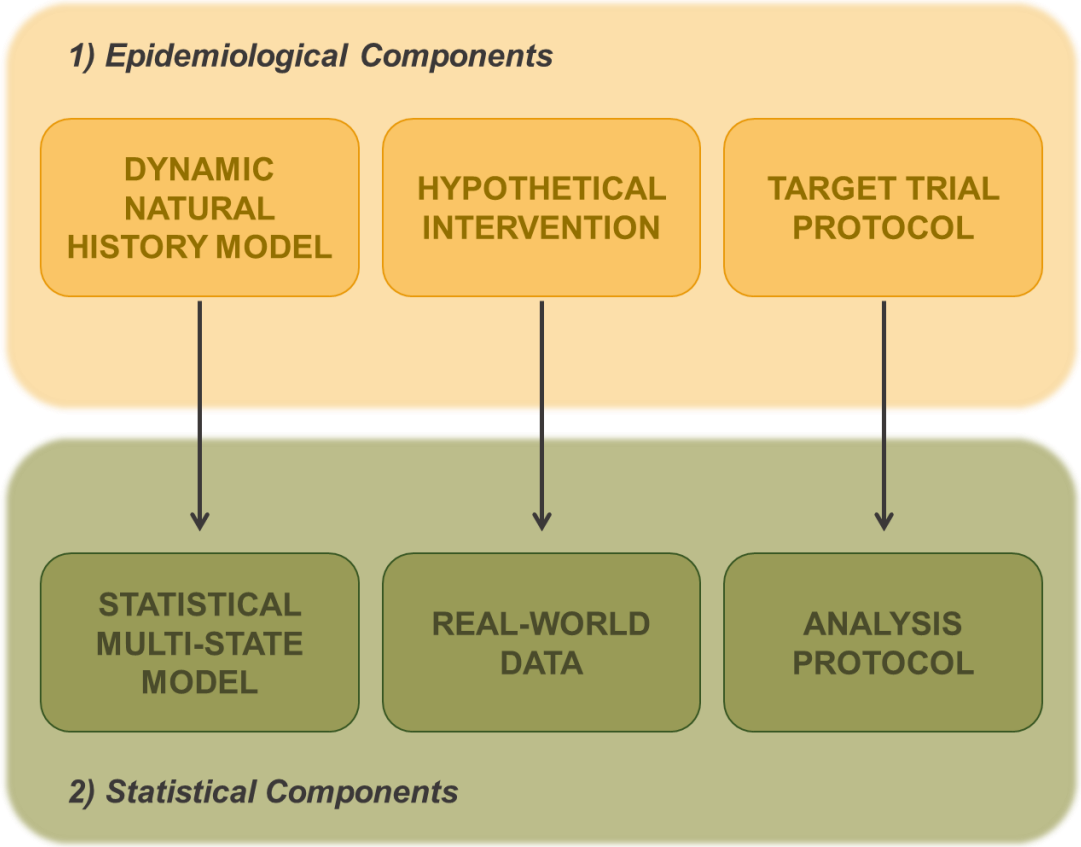


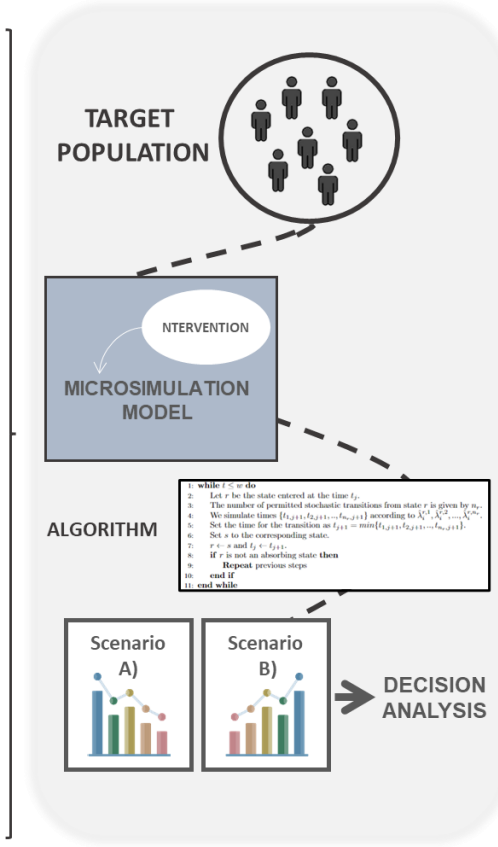
FREQUENTIST FRAMEWORK: we sample the parameters from the asymptotic normal distribution of the MLE (Parametric Bootstrap) or suitable parametric family as appropriate

BAYESIAN FRAMEWORK: we sample the parameters from their posterior probability density

Dynamic scenario analysis

Dynamic scenario analyses can be useful to answer **clinical or public health research question** which requires following the **life course trajectories of a individuals from a population across different health states.**

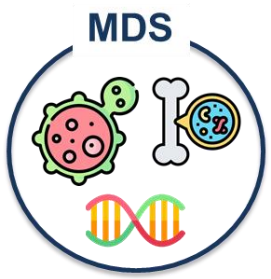




Why scenario analyses?

- Time requirements
- Target trial emulation
- Analytical difficulties
- Combine evidence from multiple sources of information

Examples of research applications of multi-state models and microsimulation



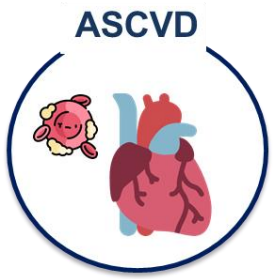
POLITECNICO MILANO 1863

HUMANITAS RESEARCH HOSPITAL

Personalized optimal timing for a treatment

Gregorio, C., Spreafico, M.,…., Della Porta,M.G., Ieva, F. **Personalized timing for allogeneic stem cell transplantation in haematological neoplasms: a target trial emulation approach using multi-state modellin and microsimulation.** JCO CCI. doi: 10.1200/CCI.23.0020

Tentori*, C.A., Gregorio, C.*,,..., R.,Ieva**, F.,Della Porta**.
Clinical and genomic-based Decision Support System to define the optimal timing of allogeneic hematopoietic stem cell transplantation in patients with myelodysplastic syndromes.* co-first; **co-last. Journal of Clinical Oncology. doi:10.1200/JCO.23.02



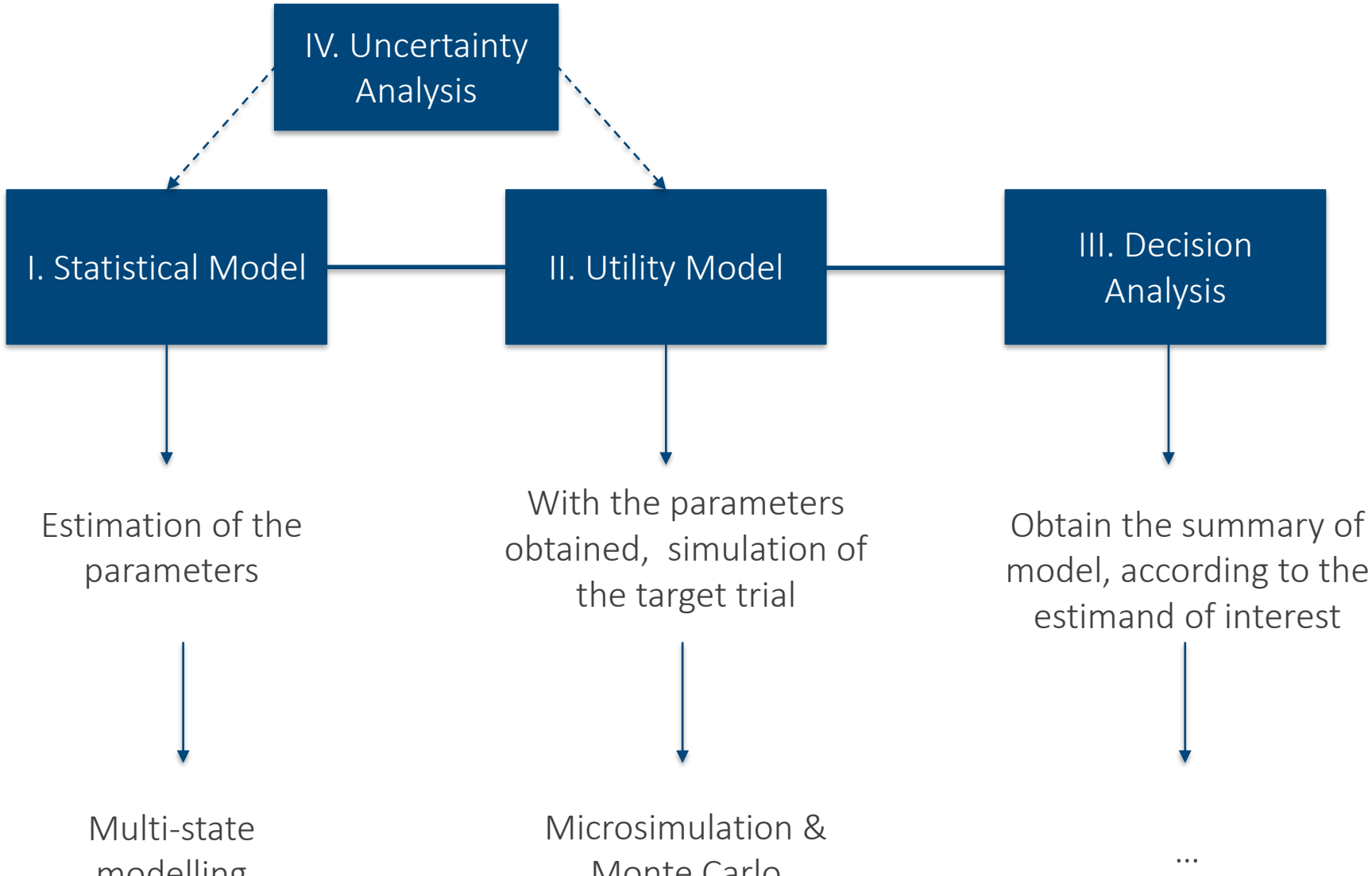
UNIVERSITÀ DEGLI STUDI DI TRIESTE

ASUGI Azienda Sanitaria Universitaria Giuliano Isontina
REGIONE AUTONOMA FRIULI VENEZIA GIULIA

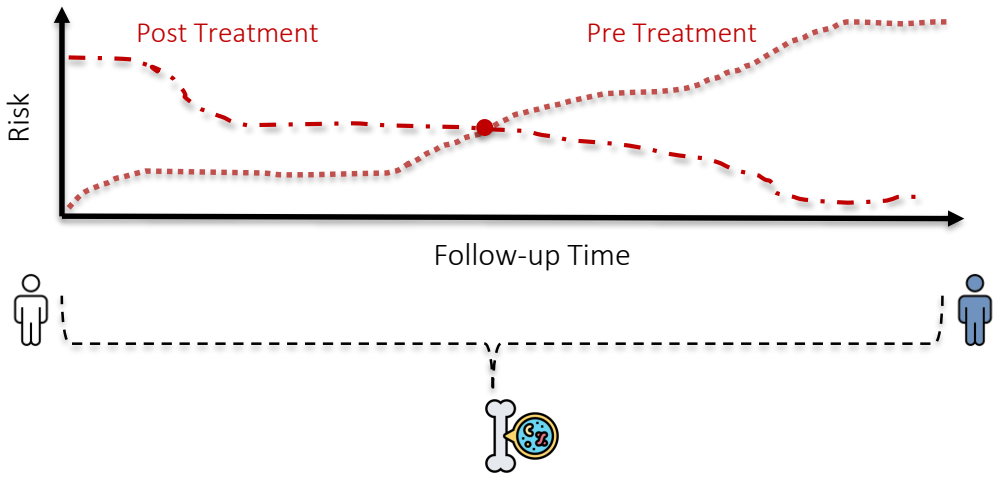
Cost-effectiveness of a treatment

Gregorio, C., Rea, F., Ieva, F., Scagnetto, A., Indennidate, C., Cappelletto, C., Di Lenarda, A., Barbati, G. **Flexible approaches based on multistate models and microsimulation to perform real-world cost-effectiveness analyses: an application to PCSK9-Inhibitors.** Value in Health. doi: 10.1016/j.jval.2024.03.008

Overview of the statistical steps



Case study: identification of a personalized treatment timing



In many clinical applications, it is of interest to study the **decision problem** of **when** to perform a **treatment** based on available patient information at baseline to **minimize** the **overall risk** of **adverse events**.

Ideally, we would like to design a randomized clinical trial but it is often impossible for practical or ethical reason.

→ **The aim of this work is to propose statistical methods to tackle this problem using available observational data**

HOMOGENEOUS EFFECT OF THE TREATMENT TIMING	vs.	ETEROGENEOUS EFFECT OF THE TREATMENT TIMING
Unique optimal timing for the whole population		Optimal timing depends on **subjects characteristics** ↓ Personalized treatment decision

Clinical Application & Data

Myelodysplastic syndromes (MDS) are a group of myeloid **neoplasms disorders** caused by poorly formed blood cells or ones that do not work properly.

- MDS condition range from **near normal life expectancy** to **acute myeloid leukaemia** with highly variable clinical courses
- **Only curative treatment**: allogeneic hematopoietic stem cell transplantation (**HSCT**)
  - Transplantation failure due to toxicity or disease relapse

Patient's characteristics

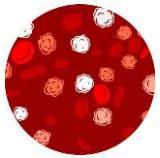
- Age at diagnosis
- Gender

Genomic information

- MDS score proposed by Bersanelli et al. 2021

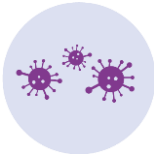
Hematological marker

- Hemoglobin [g/l]
- Neutrophils [x10^9/L]
- Platelets [x10^9/L]
- % bone marrow blasts



Risk stratification

- IPSSR score



Relapse, AML & Survival

- Relapse after HSCT
- Outcome at last follow-up

Treatment

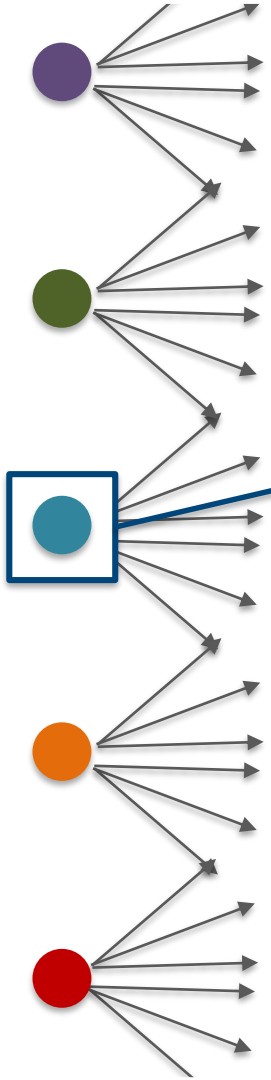
- HSCT status
- Timing

Around 7000 patients with MDS diagnosis coming from a international multi-center clinical register

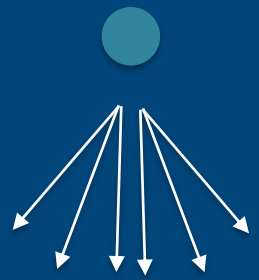
HUMANITAS
RESEARCH HOSPITAL

Target Trial

According to Hernan and Robins (2006), we imagine how a clinical randomized trial would look like in this setting



Treatment strategy allocation



Arm 1: Treat after 1 months
Arm 2: Treat after 2 months
Arm 3: Treat after 3 months
…
Arm t Treat after t months

Arms

T=1
T=2
T=3
T=4
T=5
T=…
T=…
T=…
T=…

Trial Schedule

1. Wait until 5 months
2. At 5 months perform HSCT **if patient still alive**
3. At the end of study calculate **utility function** for each subject

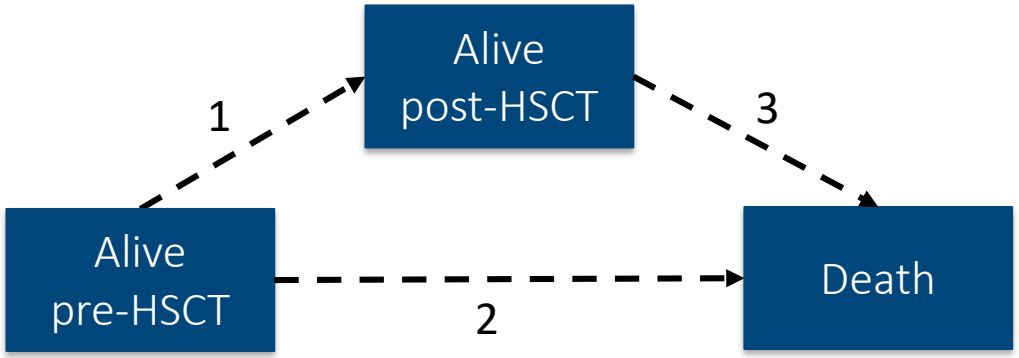
Marginal

Conditional

We want to find the timing which maximize the CEOI, conditionally on x_1

$$t^*_{x_1} = \max_{g \in [t_1; t_2]} \text{QA} - \text{RMST}^g_{x_1}$$

Multi-state model

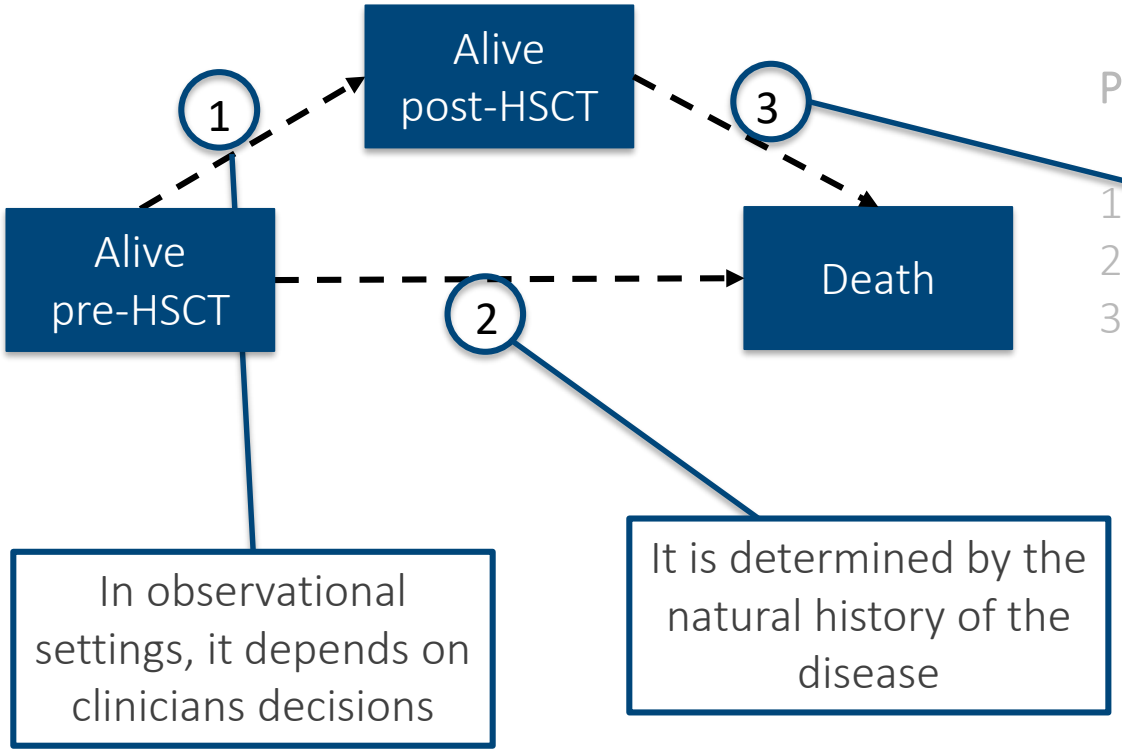


Possible states:

1. Alive pre-HSCT
2. Alive post-HSCT
3. Death

- **Transition 1:** model for the hazard of having had HSCT over time
- **Transition 2:** model for the hazard of dying before receiving HSCT over time
- **Transition 3**: model for the hazard of dying after receiving HSCT over time

Multi-state model



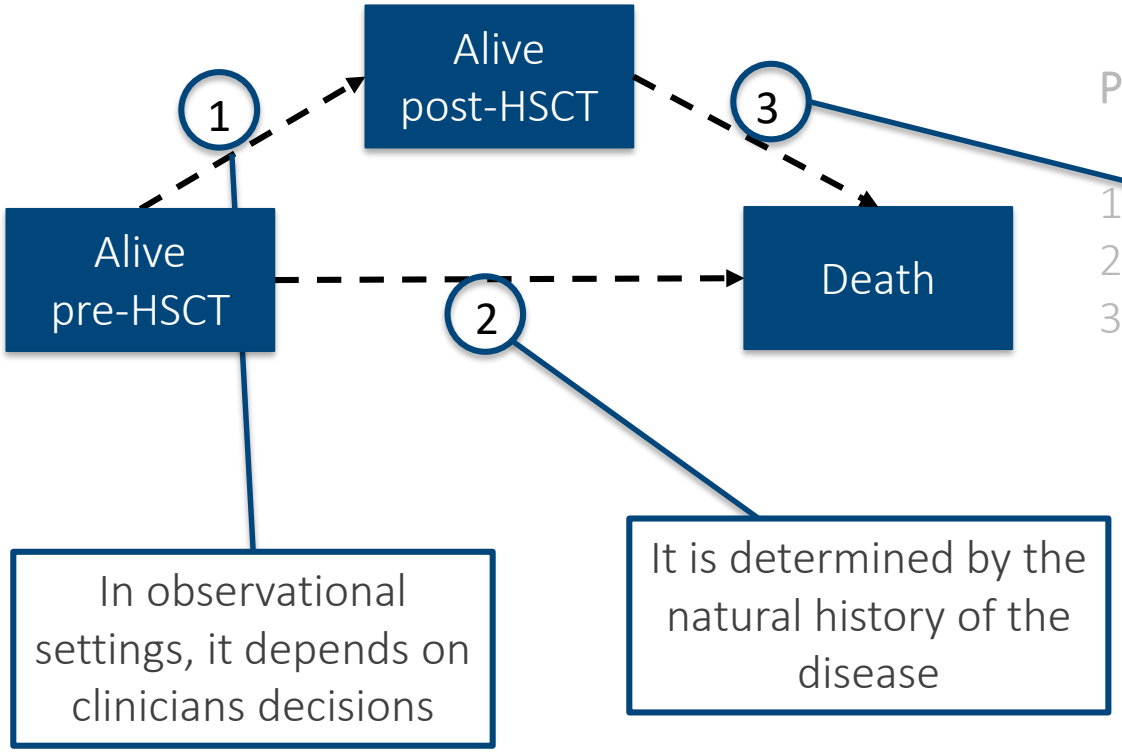
Possible states:

1. Alive pre-HSCT
2. Alive post-HSCT
3. Death

In observational settings two aspects contribute to it:

- Treatment Effect
- **Confounding because of lack of randomization**

Multi-state model



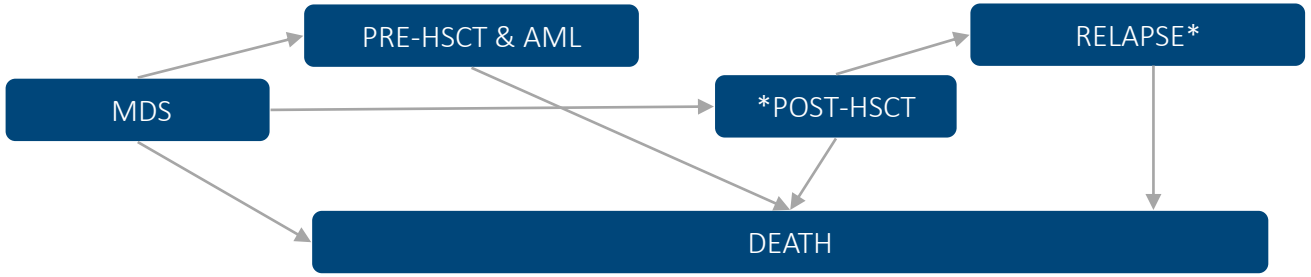
Possible states:

1. Alive pre-HSCT
2. Alive post-HSCT
3. Death

In observational settings two aspects contribute to it:

- Treatment Effect
- **Confounding because of lack of randomization**

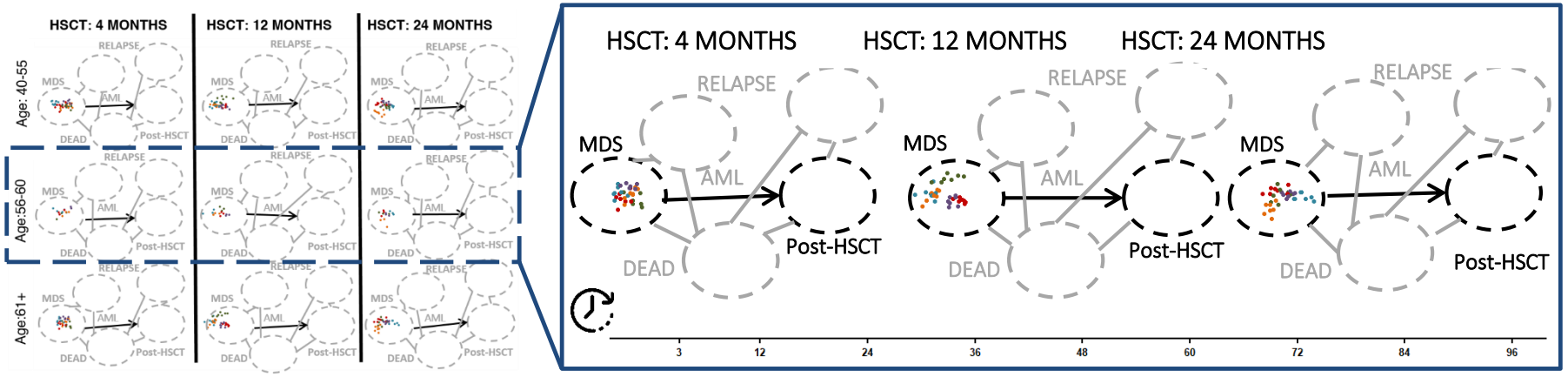
… but it is actually more complicated!



Microsimulation and decision analyses step

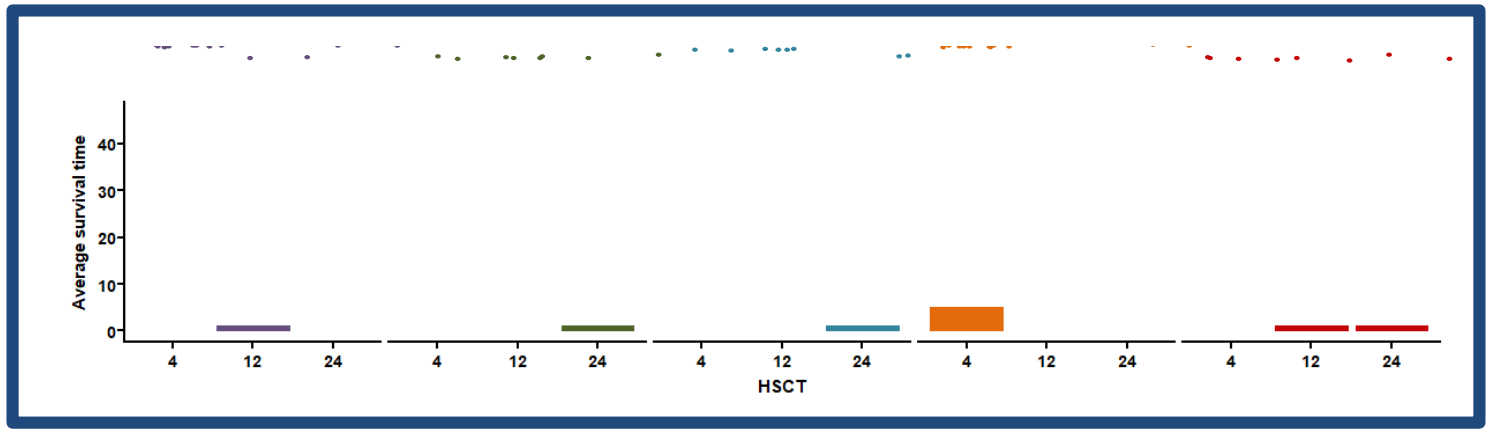
Scenario analysis

Decision Model Through Microsimulation



Decision Analysis

Optimal timing for HSCT



Case-study: cost-effectiveness of a treatment

In many clinical applications, it is of interest to study the **impact** of a given intervention taking into its **costs & benefits** in order to allocate resources optimally.

Randomized clinical trials have many limitations

→ Cost-effectiveness analyses propose statistical methods to tackle this problem using available observational data

HOMOGENEOUS EFFECT OF THE TREATMENT	vs.	ETEROGENEOUS EFFECT OF THE TREATMENT
Unique "optimal" treatment for the whole population		"Optimal" treatment depends on **subjects characteristics**
↓		↓
Is a policy cost-effective on **population** level?		Is a policy cost-effective on a **specific group of individuals** ?

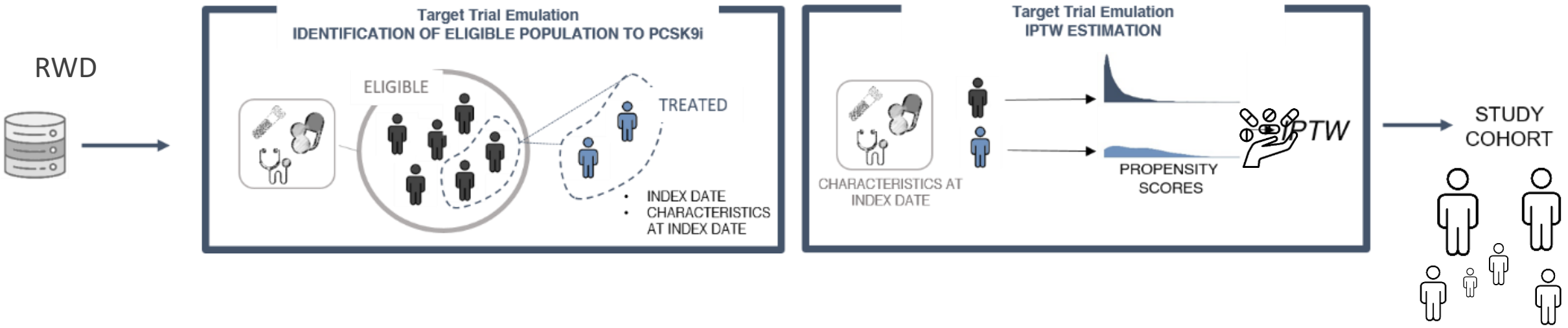
Data & clinical study

DATA SOURCE

Administrative regional health data of the Friuli Venezia Giulia region

The use of PCSK9 inhibitor drugs is governed by the decision 172/2017 of the Italian Medicines Agency (AIFA), which establishes the relative reimbursement criteria 7. This determination has recently been modified.

AIM: to evaluate cost-effectiveness of treatment with PCSk9i in eligible patients in FVG.



Statistical models in event-history settings for cost-effectiveness analysis

The statistical model are defined by three components:

- **DISEASE MODEL:** multi-state model containing all states describing events/disease stages affecting costs and benefit over time;
- **COST MODEL:** it defines the costs associated with each state, it can be either deterministic of stochastic.
- **HEALTH OUTCOME MODEL :** it defines the benefit associated with each state.

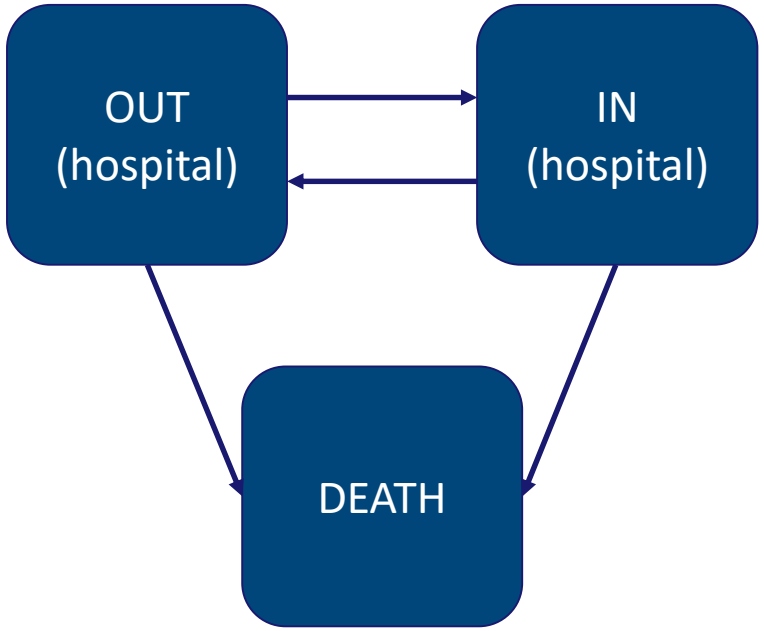
In this case:

- **DISEASE MODEL:** estimated from the data;
- **COST MODEL:** it is defined by regional price list for FVG.

We know the daily cost of being treated with statins and PCSKi9
We know the cost of hospitalizations depending of type and length

- HEALTH OUTCOME MODEL :
  - Weight fixed to 1 for all states.
  - Defined with QALY after hospitalizations

Multi-state model



We can model the transition rates using a **proportional hazard** model (Ieva et al.,2017):

$$\lambda_i^{rs}(u|\boldsymbol{x_i}) = \lambda_0^{rs}(u)\exp\{\boldsymbol{\beta_{rs}^T x_i}\}$$

Treatment and optional baseline covariates

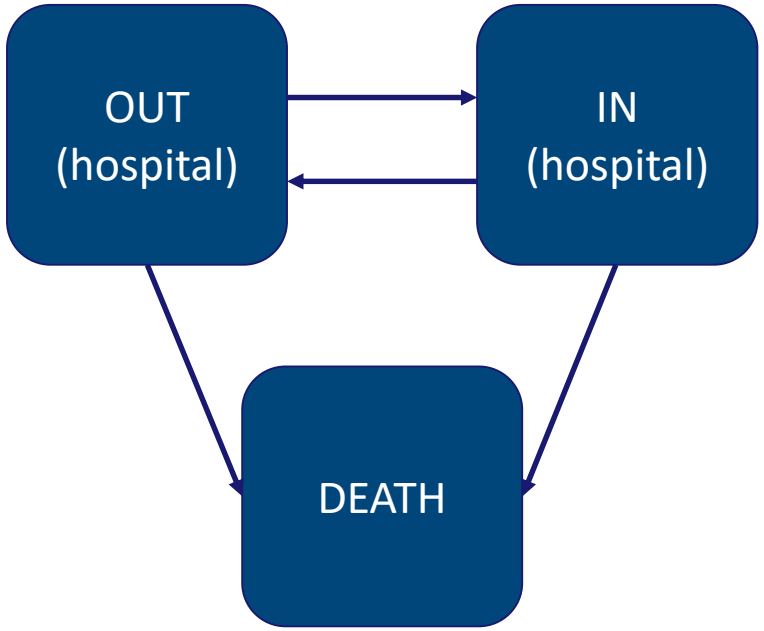
To study the:

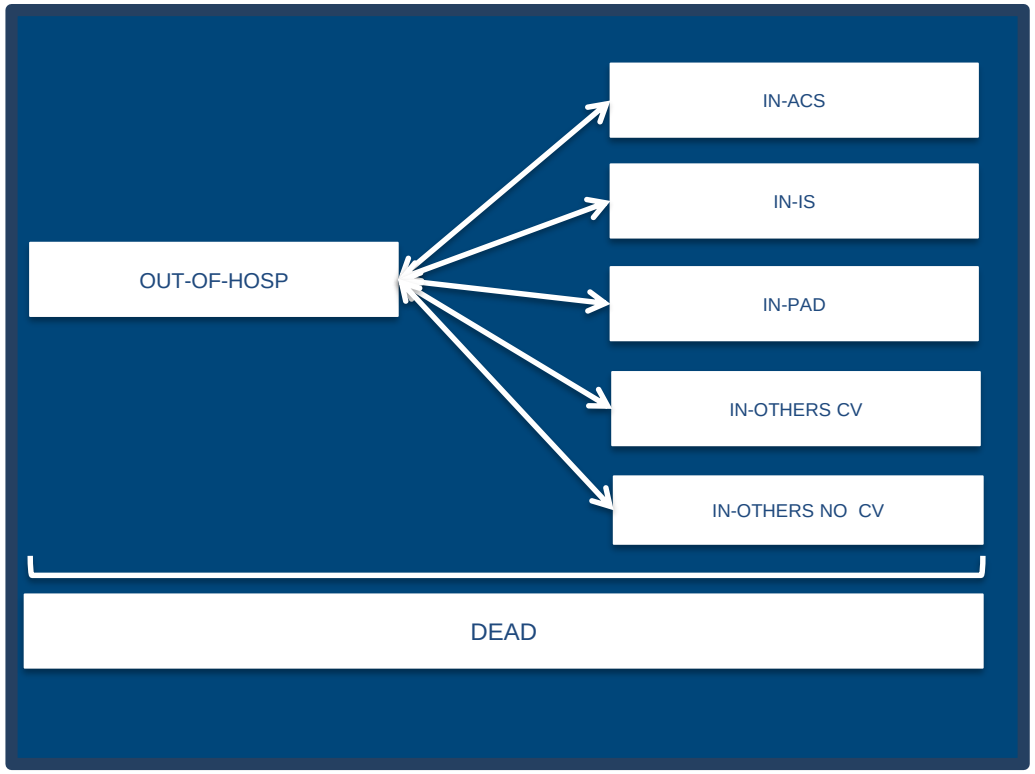
- Re-admission rate
- Discharge rate
- IN-hospital death
- OUT- of-hospital death

→ **Clock reset time scale (Semi-Markov model)**

… it is also important to take into account for **previous history**: e.g. number of past hospitalizations

Multi-state model

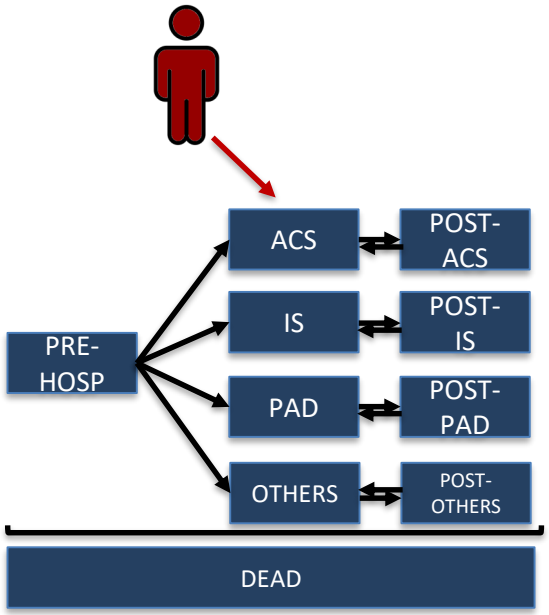




… it is also important to take into account for **previous history**: e.g. number of past hospitalizations

Cost model… an example

What is the cost attributed to a subjects spending 10 days in hospital for Acute Coronary Syndromes?



Possible Diagnosis Related Groups

557
124
558
122
121
548
550
555

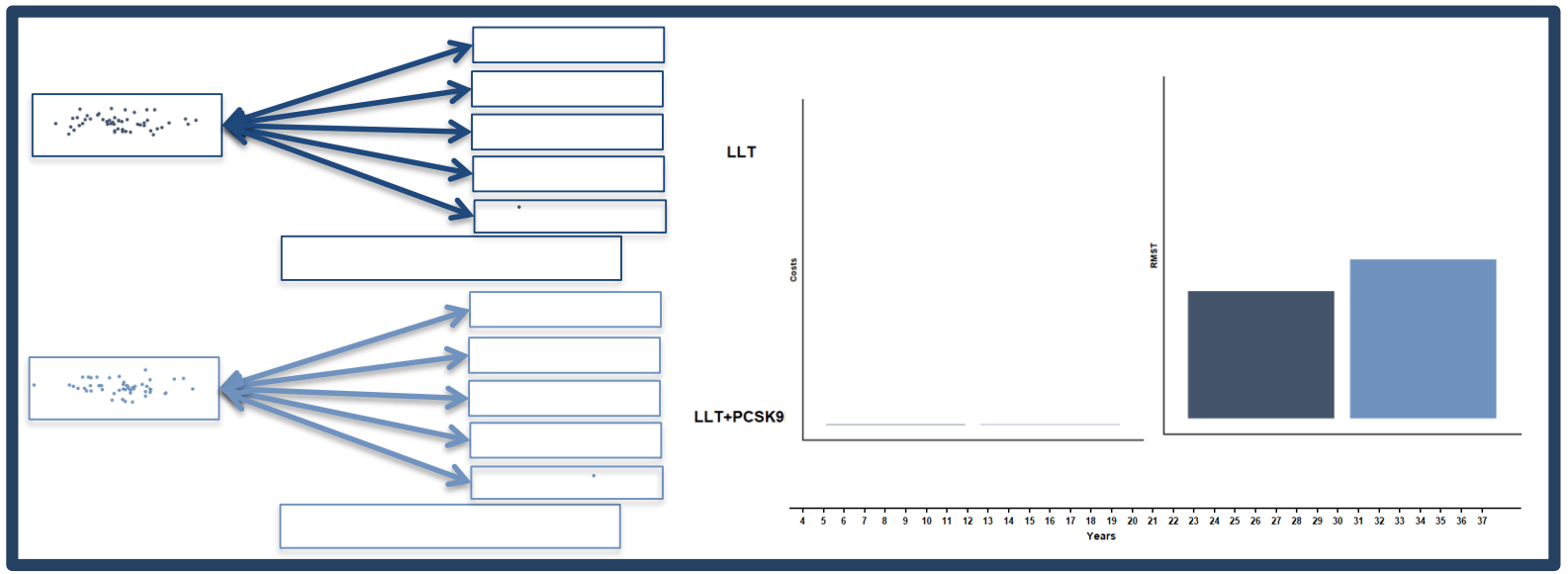
DRG 124:

Fixed cost if duration below threshold: 5.037,00 euros
Threshold: 37 days
Variable cost per extra-day: 502,00 euros

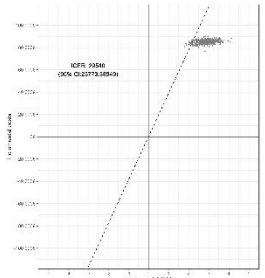
Cost rules according to the Italian national health system!

The hospitalization of patient A costed 5.037 euros!

Microsimulation and health-economic evaluation



$$ICER^W = \frac{E[C_1^W] - E[C_0^W]}{E[RMST_1^W] - E[RMST_0^W]}$$



W TIME HORIZON OF INTEREST

PCSK9i COST-EFFECTIVENESS PROFILE in REAL-WORLD

Conclusions

Given the large amount of **observational longitudinal data available**, this modeling framework provides a powerful and flexible way to try to answer a variety of questions that relates to developing and progression of **chronic diseases** over the **lifespan of individuals**.

… using models that taking into account the complexities of the phenomena under study

… from an individualized prospective

… possibly integrating causal inference methods

… allowing for simulating under “alternative scenarios” and assess the impact of possible interventions

With respect to few years ago, what makes a lot of this possible is the advance in computational power!

Thank you for the attention!

Contact information:

Caterina Gregorio (caterina.gregorio@ki.se)

References

Therneau, T.M., Grambsch, P.M. (2000) Modeling Survival Data: Extending the Cox Model. Springer, New York

Geskus, R.B., (2016). Data Analysis with Competing Risks and Intermediate States. CRC Press by Taylor and Francis Grou

Hernán, Miguel A., and James M. Robins. "Using big data to emulate a target trial when a randomized trial is not available." *American journal of epidemiology* 183.8 (2016): 758-764.

M. Skypala. Causal inference with multi-state models, applied on estimating the effect of the iui treatment timing for couples with unexplained subfertility. Master's thesis, Delft University of Technology, 2021. URL http://resolver.tudelft.nl/

uuid:37e09c1e-68fa-4419-b0f5-682b41e1f253.

Baio G, Dawid AP. Probabilistic sensitivity analysis in health economics. *Statistical methods in medical research*. 2015 Dec;24(6):615-34. doi: 10.1177/0962280211419832. Epub 2011 Sep 18. PMID: 21930515.

Ieva, F., Jackson, C.H., Sharples, L.D. (2017) Multi-State modelling of repeated hospitalisationand death in patients with Heart Failure: the use of large administrative databases in clinical epidemiology. Statistical Methods in Medical Research, 26 (3):1350-1372

Gregorio C. Toward a dynamic approach to decision-making in chronic illness : a statistical framework using real-world longitudinal data. PhD Thesis. https://hdl.handle.net/10589/216672

R packages:

Jackson, C., & Jackson, M. C. (2021). Package 'flexsurv'. *Flexible parametric survival and multi-state models, https://cran. r-project. org/web/packages/flexsurv/flexsurv. pdf (2018, accessed 26 September 2019)*.

Bayesian Methods in Health Economics. Baio et al (2017). Bayesian Cost-Effectiveness Analysis with the R package BCEA

Incerti D, Jansen J (2021). "hesim: Health Economic Simulation Modeling and Decision Analysis." 2102.09437, https://arxiv.org/abs/2102.09437.

de Wreede LC, Fiocco M, Putter H (2011). "mstate: An R Package for the Analysis of Competing Risks and Multi-State Models." Journal of Statistical Software, 38(7), 1–30. doi:10.18637/jss.v038.i07, https://www.jstatsoft.org/v38/i07/.

de Wreede LC, Fiocco M, Putter H (2010). "The mstate Package for Estimation and Prediction in Non- and Semi-Parametric Multi-State and Competing Risks Models." Computer Methods and Programs in Biomedicine, 99, 261–274.

Putter H, Geskus RB, Fiocco M (2007). "Tutorial in Biostatistics: Competing Risks and Multi-State Models." Statistics in Medicine, 2389–2430.

Therneau T (2023). *A Package for Survival Analysis in R*. R package version 3.5-5, https://CRAN.R-project.org/package=survival.