



# Modeling the cumulative incidence function of clustered competing risk data

Henrique Laureano

<http://leg.ufpr.br/~henrique>

Laboratório de Estatística e Geoinformação (LEG)  
Programa de Pós-Graduação em Métodos Numéricos em Engenharia (PPGMNE)  
Universidade Federal do Paraná (UFPR)

December/2020



## Clustered competing risk data

Key terms:

- » **Clustered**: groups with a dependence structure (e.g. families);
- » Causes **competing** by *something*.

Something?

- » **Failure** of an industrial or electronic component;
- » **Occurrence** or **cure** of a disease;
- » **Progress** of a patient clinic state.

Independent of the application, always the same framework

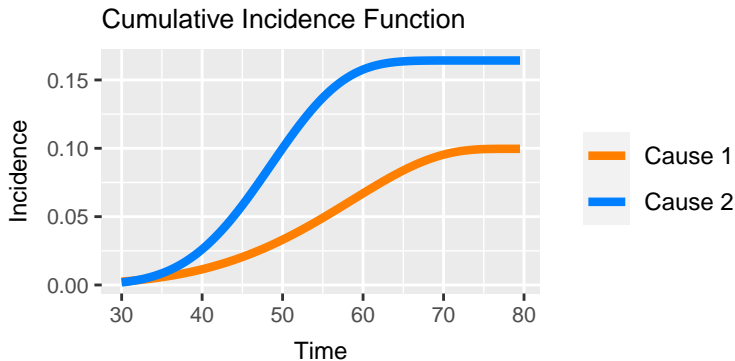
Cluster	ID	Cause 1	Cause 2	Censorship	Time
1	1	Yes	No	No	10
1	2	No	No	Yes	8
2	1	No	No	Yes	7
2	2	No	Yes	No	5



## What we do?

We model the **probability** of each **competing cause** along the time and taking into account the possible **within-cluster dependence**

... all this in terms of a



Main focus application:  
cancer incidence in twins



Clustered competing risks data

↳ Clusters? Families

↳ Family studies

↳ **Twins data**

Family studies  $\Rightarrow$  **within-family dependence**

That may reflect

- » Disease **heritability**;
- » The impact of shared **environmental effects**;
  - » **Parental effects**:



## Challenges

- » The data is very simple, we just know the outcome (**yes** or **no**);
- » With this, we have to be able to construct the **cumulative incidence curves**;
- » And we have to accommodate the **within-family dependence** properly, that can happen in different manners;
- » ...



Thank you



Joint work with

Wagner H. Bonat

<http://leg.ufpr.br/~wagner>

Paulo Justiniano Ribeiro Jr.

<http://leg.ufpr.br/~paulojus>



@hap\_laureano

