ISCB43
NEWCASTLE, UK | 2022

Subgroups identification with survival decision trees: detection of early / late conversion between Alzheimer's stages

Marie Laurent*[1], Martin Prodel[1], Alexandre Vainchtock[1], Alexandre Civet[2], Antoine Chevrette[3], Béatrice Rousselle[2], Claire Moutet[4], Virginie Dauphinot[4], and Pierre Krolak-Salmon[4]

[1] **Heva**, Lyon, France
[2] **Roche France**, Boulogne Billancourt, France
[3] **Roche Global**, Basel, Switzerland
[4] Centre Mémoire de Ressources et de Recherche de Lyon, **Hospices civils de Lyon**

ML, MP and AV are employees at HEVA.
ACI, ACH and BR are employees at ROCHE.
CM, VD and PKS are employees at Hospices Civils de Lyon (HCL).
This study was financed by Roche France and implemented by HEVA and the HCL.

The MEMORA database

Data collected in memory centers of the Clinical and Memory Research Center of Lyon
Objective of the cohort: to investigate the factors associated with functional autonomy change and with the cognitive performance and behavioral disorders changes over time.

Demographic information

Available

- ✓ Age
- ✓ Gender
- ✓ Educational level
- ✓ Profession (when the patient was working)

Medical information

Available

- ✓ Treatments
- ✓ Comorbidities [1]

Claim hospital data (Diagnoses)

- ○ (for 62 % of patients)

Scores specific to neurodegenerative diseases

Available

- ✓ MMSE (memory)

Variables with more than 20% of missing values were not included in the analyses
(Ex: IADL, Grober and NPI scores, lumbar puncture results)

[1] identified through treatments (text field), comorbidities, antecedents (text field), and diagnoses in the hospital claim database

Cohort et stages

Definitions

Definition of the cohort

Inclusion period: 2014 to 2019

5 210 Alzheimer's patients

↓ Exclusion on age (< 50 or > 90)

4 672 patients

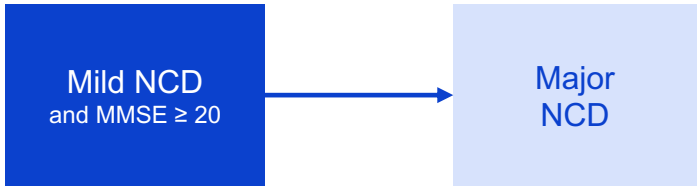
↓ Exclusion on comorbidities [1]

4 028 patients

↓ Selection of patients with mild NCD

1,264 patients with mild NCD and MMSE ≥ 20

Definition of the disease stages



Diagnostic **Mild or Major** set by a clinician

MMSE score « Mini Mental state examination »
30-item questionnaire assessing patient's memory and concentration (30: no dementia, 0: dementia)

[1] Parkinson, Lewy body dementia, Frontotemporal degeneration, Glioma, Meningioma, Multiple sclerosis, Lupus, Antiphospholipid, Behcet syndrome, HIV

Definition of a censored time to conversion

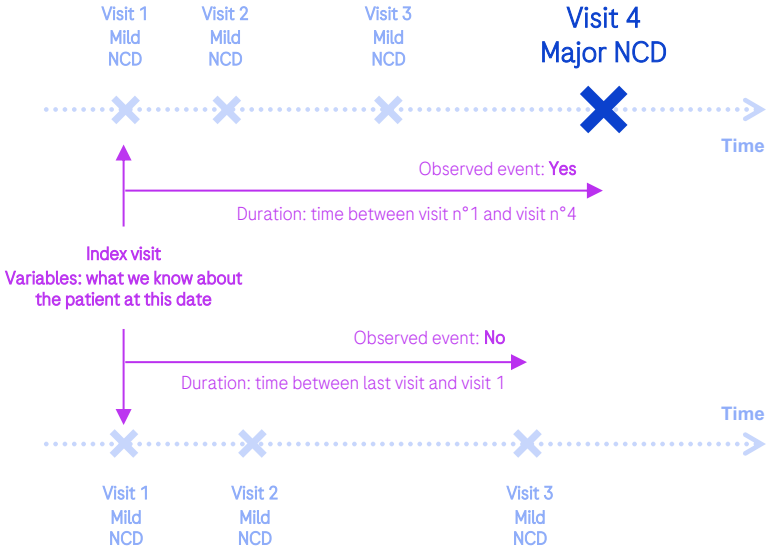
Modeling patients

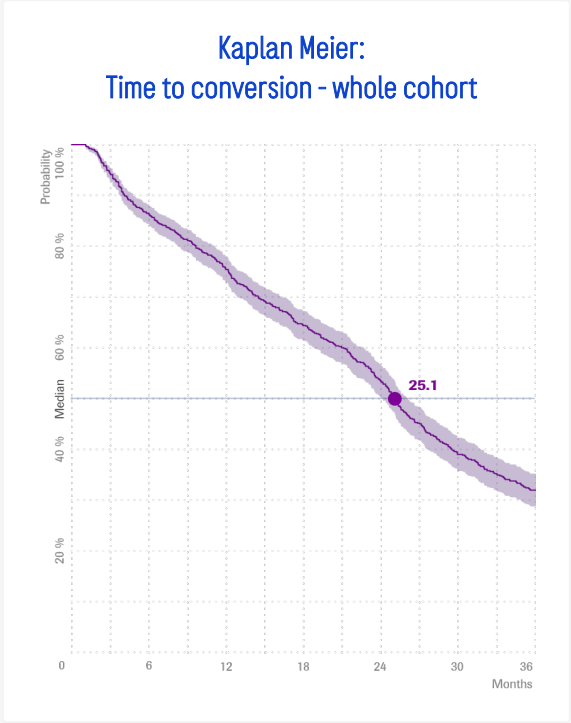
1st case: CONVERSION

The patient converts during the study period

2nd case: CENSORING

The patient does not convert during the study period





Descriptive analyses

Distribution of the 190 features extracted from the database

1,264 patients

with Mild NCD

(and MMSE ≥ 20)

Mean age

77 y.o.

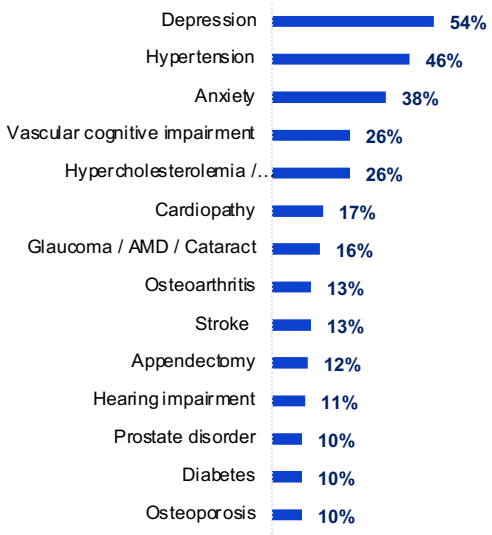
± 9 y.o.

Gender

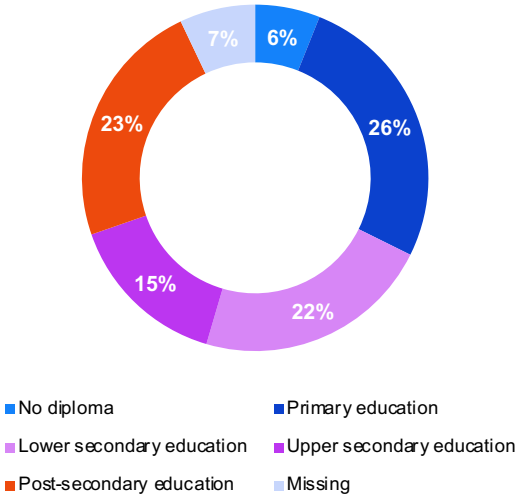
65%

of women

Comorbidities (% of patients)



Educational level

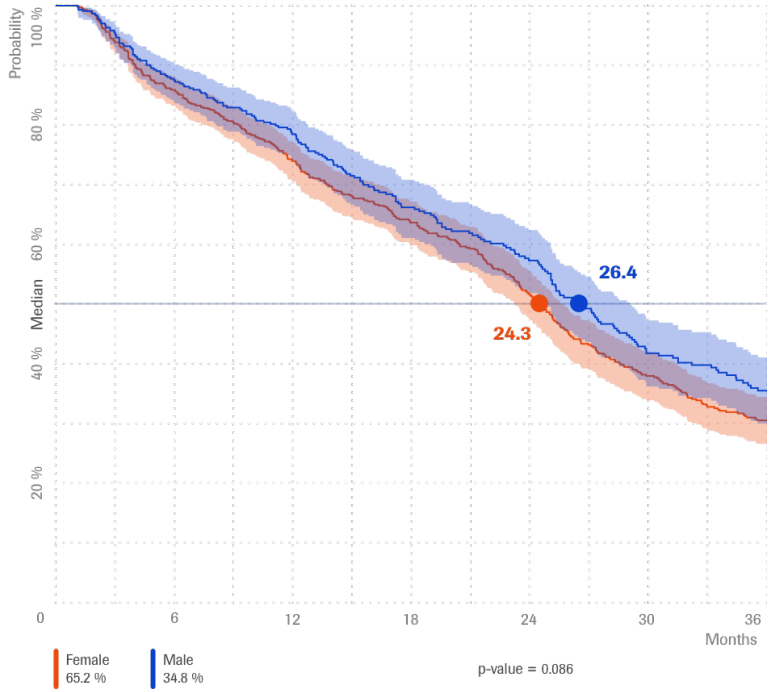


Presentation of other distributions in an interactive dashboard to simplify the collaboration with medical experts

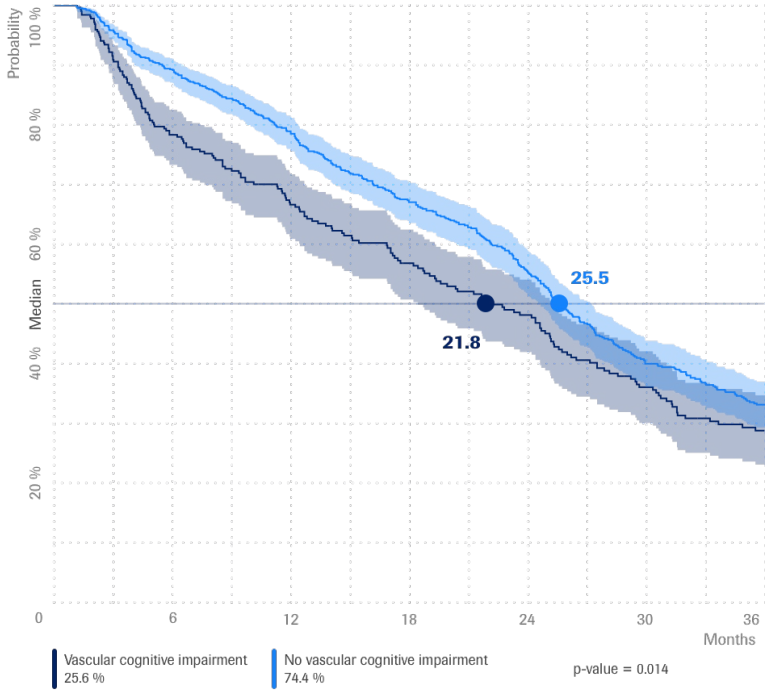
Univariate analysis

Kaplan Meier of the time to conversion with regard to each variable

Gender



Vascular cognitive impairment



Presentation of other Kaplan Meier conversion times in an interactive dashboard

Method: the survival tree

A mix between decision tree and survival analysis

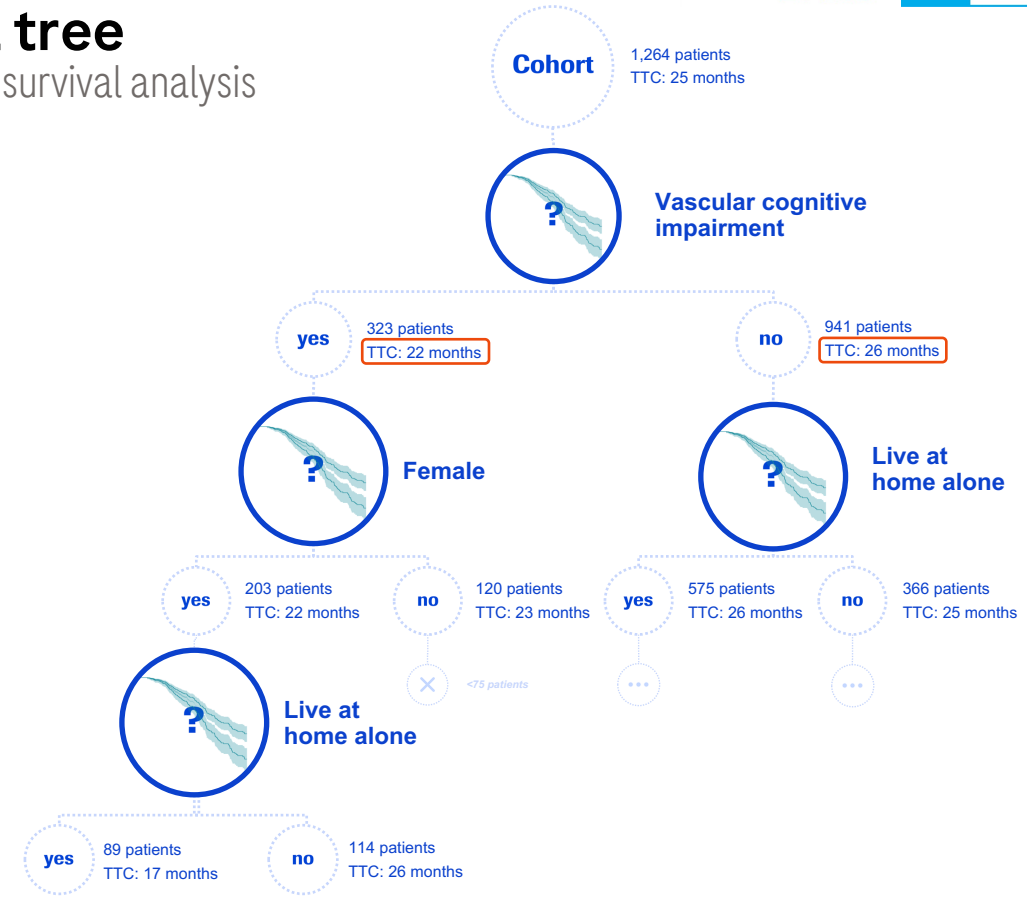
Decision tree (machine learning)

« Find subgroups of patients different …

+

Survival analysis

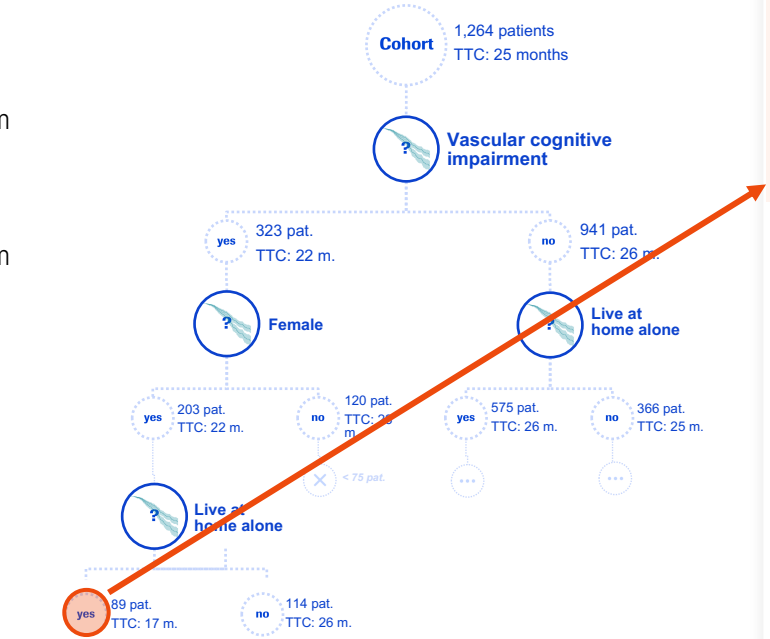
… in the way that their censored time to conversion is significantly higher or lower »



Method: subgroup selection among nodes

Conditions for a node to be selected:

- ✓ % patients > 6%
- ✓ Subgroup's TTC different from its complementary in the cohort *(p-value < 0.1)*
- ✓ Subgroup's TTC different from its sibling *(p-value < 0.1)*
- ✓ Subgroup's median TTC more distant than 15% of cohort's median TTC
- ✓ Subgroup's median TTC more distant from cohort's median TTC than sub-group parent's median TTC

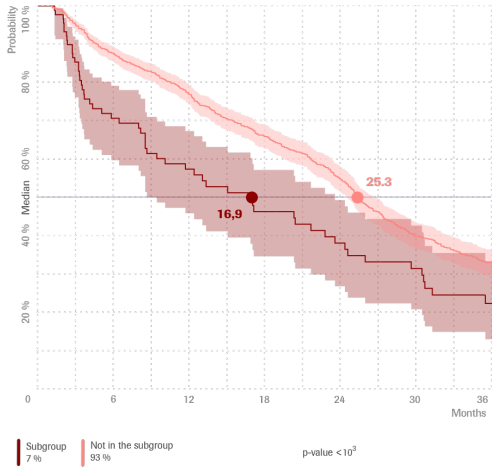


selected node

A subgroup with an early conversion:

Median time to conversion: **16.9** months
89 patients (7%)

Characteristics:
1. Vascular cognitive impairment
2. Gender: Female
3. Live at home alone

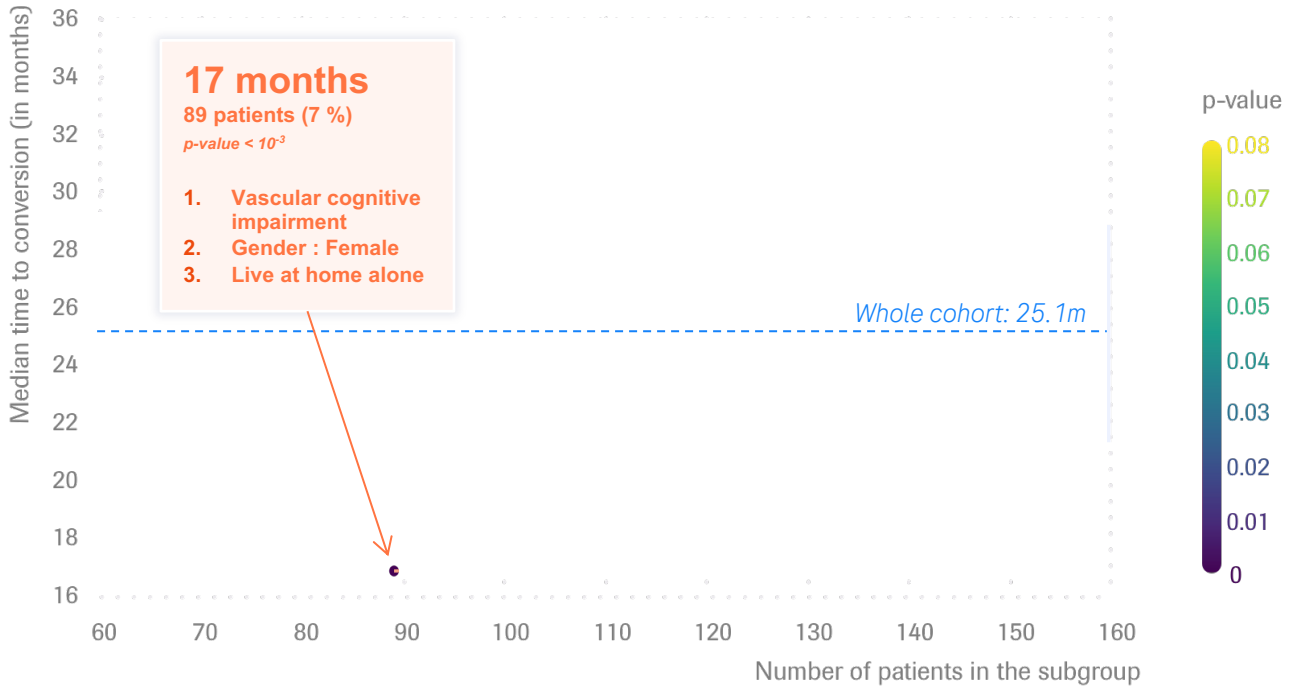


Method: selection of relevant subgroups

Representation of tree nodes regarding the number of patients and the median time to conversion

Bagging:
Train several survival trees with different subsets of features

→ Obtain a variety of subgroups



Method: selection of relevant subgroups

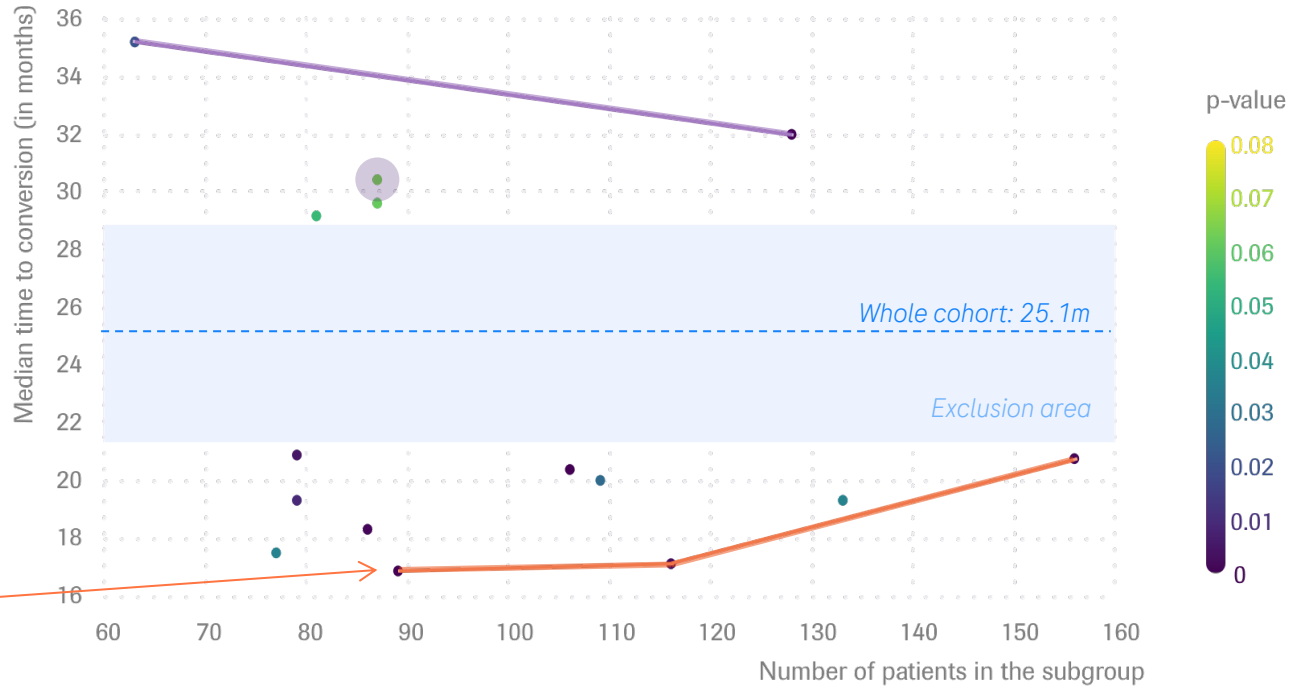
Representation of tree nodes regarding the number of patients and the median time to conversion

Bagging:
Train several survival trees with different subsets of features

→ Obtain a variety of subgroups

17 months
89 patients (7 %)
p-value < 10^{-3}

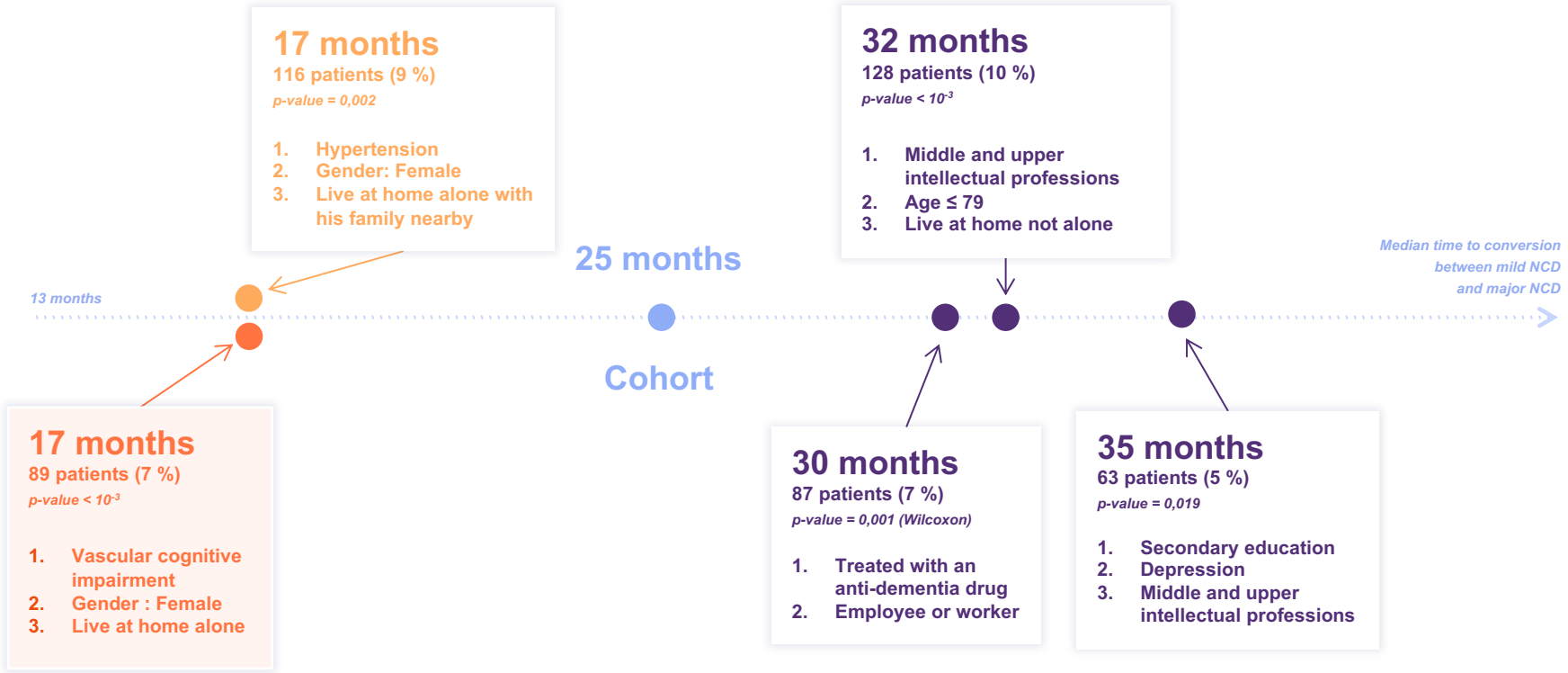
1. **Vascular cognitive impairment**
2. **Gender : Female**
3. **Live at home alone**



Use of the Pareto fronts to select profiles

Results

6 subgroups of patients converting early / late from mild NCD to major NCD stages



Conclusion

- The **MEMORA database** is a source of real-life information on Alzheimer's patients.
- Patients with a **higher education** converted later, as well as patients who took a **symptomatic drug for Alzheimer's disease**. Whereas patients with **vascular cognitive impairment** converted earlier (univariate analysis).
- **Survival decision trees** helped us to identify subgroups of patients converting early or late from mild to major NCD stages. While the **decision tree** identifies **combinations of features** impacting the conversion time (whereas a Cox model studies the impact of each feature one by one), the **survival** allowed us to model a **censored target**.
- For example, **living at home alone** has a low impact on the entire cohort, but a big impact on the time to conversion **among women with a primary education or no diploma.**

Thank you for your attention

Martin Prodel : mprodel@hevaweb.com