Visualization methods combined to clustering approach enabled the identification of clinically relevant patterns of prostate cancer management.

ABSTRACT

BACKGROUND

There is a great heterogeneity of prostate cancer patient journeys for a same stage of disease, which makes their evaluation complex using descriptive statistics. Unsupervised machine learning has the potential to reveal patterns within heterogeneous data. We aimed to illustrate how clustering and visualization of healthcare pathways can enhance the characterization of patients with Prostate Cancer (PC), at all disease stages.

METHODS

Data source & Tools:

- French nationwide healthcare database (*Système National des Données de Santé*, SNDS).
- R packages used: clustofVar, TraMineR, WeightedCluster

Study population & Data:

- Participants: Patients with prevalent prostate cancer in 2014 were identified among men alive, aged ≥40, and covered by the general health insurance scheme in 2014 based on discharge diagnosis and specific encounters.
- Observation period: Data were extracted with a 5-year history period, and with up to 7 semesters of follow-up from the PC diagnosis.
- A validated algorithm* was applied to detect castration resistance, metastasis management and so metastatic and castration-resistant status.
- Four exclusive **cohorts of incident patients** were constituted, prioritizing the most advance disease stage:
  - ✓ hormone-sensitive (**HSPC**),
  - ✓ metastatic hormone-sensitive (**mHSPC**),
  - ✓ castration-resistant (**CRPC**),
  - ✓ metastatic castration-resistant (**mCRPC**).

Study stages and Clustering method:
See beside.

RESULTS

Clustering allowed to distinguish trends in healthcare pathways such as patients:

- Undergoing local treatment with or without androgen deprivation therapy in HSPC,
- With rapid and slow disease progression in CRPC,
- With curative, or palliative intent treatment in mHSPC and mCRPC.

CONCLUSION

Visualization methods combined to clustering approach enabled the identification of clinically relevant patterns of prostate cancer management. Characterization of these care pathways is an essential element for the comprehension and the robust assessment of healthcare technologies effectiveness.

*Thurin NH, Rouyer M, Gross-Goupil M, et al. Epidemiology of metastatic castration-resistant prostate cancer: A first estimate of incidence and prevalence using the French nationwide healthcare database. Cancer Epidemiol. 2020;69:101833.

Disclosure:
The authors received no financial support for the research. However, the example used was drawn from the TherapeutiC strategy in Metastatic castration-Resistant pRostate cAncer: target population and changes between 2012 and 2014 (CAMERRA) study, which was funded by Janssen-Cilag, France and carried out by the Bordeaux PharmacoEpi platform under the supervision of a Scientific Committee.

Applying sequence clustering methods to characterize healthcare pathways of patients in the French nationwide healthcare database: Prostate cancer example

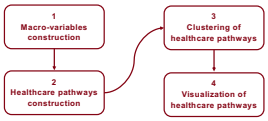
NH. Thurin[1], R. Lassalle[1], R. Baulain[2], J. Jové[1], D. Sakr[1], M. Gross-Goupil[3], M. Rouyer[1], P. Blin[1], C. Droz-Perroteau[1]

[1] *Univ. Bordeaux, INSERM CIC-P 1401, Bordeaux PharmacoEpi, Bordeaux, France*
[2] *Ecole nationale de la statistique et de l'administration économique Paris (ENSAE), Institut Polytechnique Paris, Palaiseau, France*
[3] *Medical Oncology, Hôpital Saint-André, CHU Bordeaux, Bordeaux, France*

Study stages & Clustering method

51 specific healthcare encounters constitutive of PC management were synthetized into **4 macro-variables: surveillance, local treatment, androgenic deprivation, and advanced treatment** based on the whole patient data.

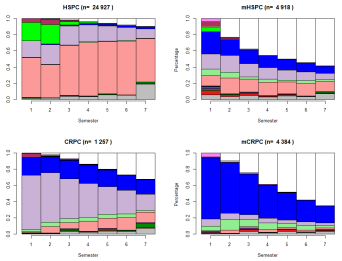
For each patient, 1 **state** was assigned per semester among the 16 possible ones according to the values of macro-variables. The succession of these states over time formed the **healthcare pathway**.



In each cohort (HSPC, mHSPC, CRPC, mCRPC) **optimal matching** (**TRATE** substitution method) was applied to calculate distances between pathways, and **Partitioning Around Medoids** (PAM) algorithm was used to generate consistent groups of similar pathways.

State Sequence Analysis (SSA) method was used to visualize clustered healthcare pathways. Optimal number of clusters was selected from internal validation indexes (Hubert coefficient, average silhouette width…).

Healthcare states



A total of 35 486 incident prostate cancer cases were identified in 2014, forming four cohorts (Figure 1):

- ✓ HSPC incident cohort
- ✓ mHSPC incident cohort
- ✓ CRPC incident cohort
- ✓ mCRPC incident cohort

For each incident cohort, Figure 1 shows the distribution of healthcare pathways states by semester.

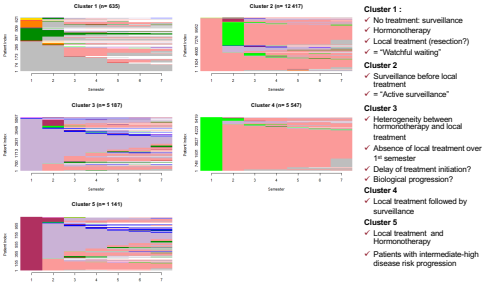
Healthcare states' Key

- no treatment
- advanced treatment (advanced treat.)
- androgenic deprivation (andro. depr.)
- andro. depr. + advanced treat.
- local treatment (local treat.)
- local treat. + advanced treat.
- local treat. + andro. depr.
- local treat. + andro. depr. + advanced treat.
- surveillance + advanced treat.
- surveillance + andro. depr.
- surveillance + andro. depr. + advanced treat.
- surveillance + local treat.
- surveillance + local treat. + advanced treat.
- surveillance + local treat. + andro. depr.
- surveillance + local treat. + andro. depr. + advanced treat.
- death

Figure 1. Distribution of healthcare states by semester among the four incident cohorts

Patients pathway

HSPC



Cluster 1 :
- ✓ No treatment: surveillance
- ✓ Hormonotherapy
- ✓ Local treatment (resection?)
- ✓ = "Watchful waiting"

Cluster 2
- ✓ Surveillance before local treatment
- ✓ = "Active surveillance"

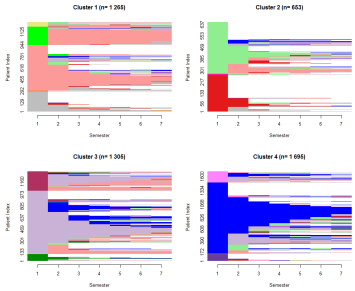
Cluster 3
- ✓ Heterogeneity between hormonotherapy and local treatment
- ✓ Absence of local treatment over 1st semester
- ✓ Delay of treatment initiation?
- ✓ Biological progression?

Cluster 4
- ✓ Local treatment followed by surveillance

Cluster 5
- ✓ Local treatment and Hormonotherapy
- ✓ Patients with intermediate-high disease risk progression

Figure 2. Cluster index plots of the incident HSPC cohort (n = 24 927)

mHSPC



Cluster 1
- ✓ Local treatment (node resection, salvage surgery with curative intent treatment?) followed by surveillance

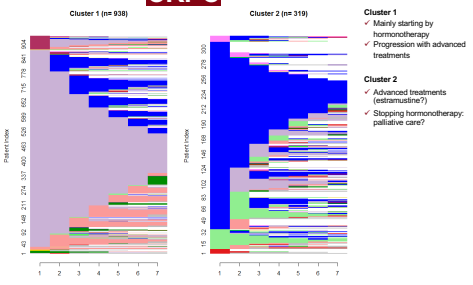
Cluster 2
- ✓ No hormonotherapy
- ✓ Low survival
- ✓ Advanced treatment (palliative intent treatment)

Cluster 3
- ✓ Hormonotherapy
- ✓ Local treatment (node resection, salvage surgery with curative intent treatment?)
- ✓ Initiation of advanced treatments

Cluster 4
- ✓ Hormonotherapy
- ✓ Initiation of advanced treatments

Figure 3. Cluster index plots of the incident mHSPC cohort (n = 4 918)

CRPC

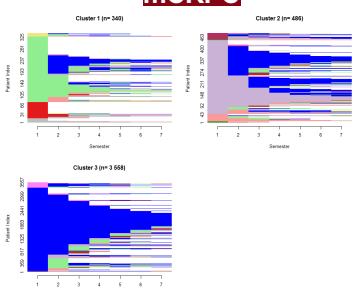


Cluster 1
- ✓ Mainly starting by hormonotherapy
- ✓ Progression with advanced treatments

Cluster 2
- ✓ Advanced treatments (estramustine?)
- ✓ Stopping hormonotherapy: palliative care?

Figure 4. Cluster index plots of the incident CRPC cohort (n = 1 257)

mCRPC



Cluster 1
- ✓ No hormonotherapy
- ✓ Advanced treatments (local treatment of metastases?)
- ✓ Palliative care?
- ✓ Low survival

Cluster 2
- ✓ Hormonotherapy
- ✓ Delayed initiation of advanced treatment
- ✓ Slow progression or delay in initiation?

Cluster 3
- ✓ Advanced treatment
- ✓ Maintained hormonotherapy

Figure 5. Cluster index plots of the incident mCRPC cohort (n = 4 384)

38th International Conference on Pharmacoepidemiology and Therapeutic Risk Management Copenhagen, Denmark, August 24-28, 2022

www.bordeauxpharmacoepi.eu
nicolas.thurin@u-bordeaux.fr
@BxPharmacoEpi

Scan to download the poster