

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



ABSTRACT

BACKGROUND

The heterogeneity of prostate cancer (PC) patient journeys for a same stage of disease makes their evaluation complex using descriptive statistics. Unsupervised machine learning has the potential to reveal patterns within heterogeneous data. However, impact of such methods in real-world studies is not yet clear.

OBJECTIVE

To illustrate how clustering and visualization of healthcare pathways can enhance the characterization of patients with 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.* 2021;68:101633.

Disclaimer

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 at different prostate cancer stages in the French nationwide healthcare database (SNDS)

NH. Thurin¹, R. Lassalle¹, R. Baulain², J. Jové¹, D. Sakr¹, M. Gross-Goupil³, M. Rouyer¹, P. Blin¹, C. Droz-Perroteau¹



¹ Univ. Bordeaux, INSERM CIC-P 1401, Bordeaux Pharmacoepi, Bordeaux, France

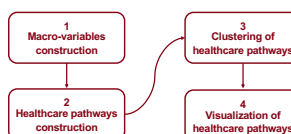
² Ecole nationale de la statistique et de l'administration économique Paris (ENSAE), Institut Polytechnique Paris, Palaiseau, France

³ Medical Oncology, Hôpital Saint-André, CHU Bordeaux, Bordeaux, France

Study stages & Clustering method

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

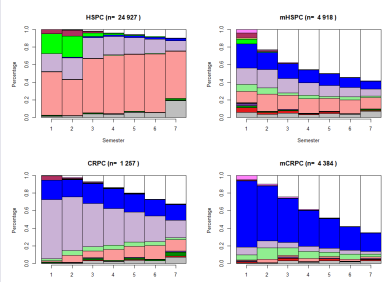
For each patient, 1 state was assigned per semester among the 16 possible ones according to the values of macro-variables. The succession of these states over time formed the healthcare pathway.



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
- 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

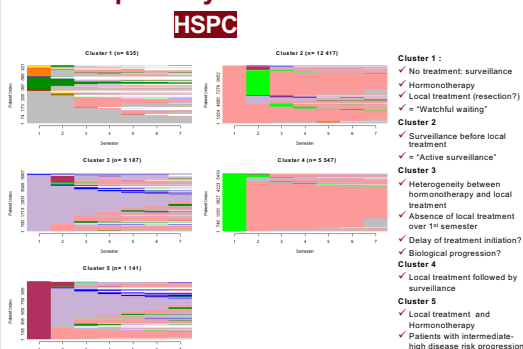


Figure 2. Cluster index plots of the incident HSPC cohort (n = 24 927)

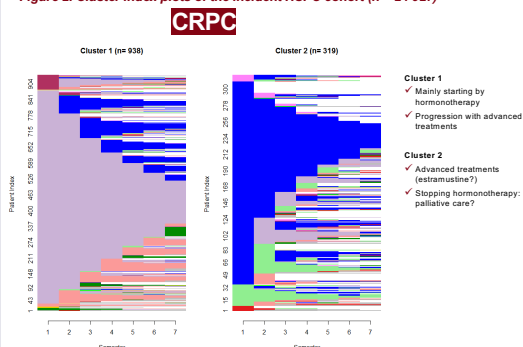


Figure 4. Cluster index plots of the incident CRPC cohort (n = 1 257)

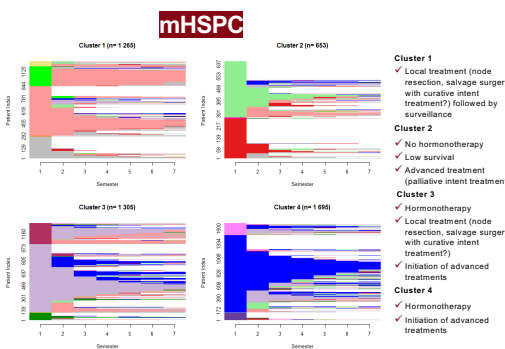


Figure 3. Cluster index plots of the incident mHSPC cohort (n = 4 918)

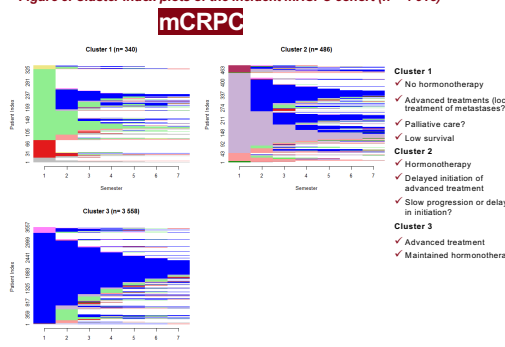


Figure 5. Cluster index plots of the incident mCRPC cohort (n = 4 384)

ISPOR Europe 2022, Vienna, Austria, November 6-9, 2022