Cluster Algorithm for Credible Subgroup Identification in Patients With Erectile Dysfunction Receiving Tadalafil: A Real-World Data Study

Kevin T. McVary[1], Chathuri Daluwatte[2], Robert A. Kloner[3], Martin Montmerle[4], Mélissa Rollot[4], Pauline Guilmin[4], Margot Blanchon[4]*, Nicolas Lambert[4], Cyril Esnault[4], Andrew Stewart[2], Thomas McGraw[5]

[1]Department of Urology, Stritch School of Medicine, Loyola University Medical Center, Maywood, IL, USA; [2]Digital Data, Sanofi Inc., Cambridge, MA, USA; [3]Cardiovascular Research, Huntington Medical Research Institutes, Pasadena, CA, USA, Keck School of Medicine, Dept of Medicine, Cardiovascular Division, University of Southern California, Los Angeles, CA, USA; [4]Quinten Health, Paris, France; [5]Sanofi Inc., CHC Science Hub, Bridgewater, NJ, USA
*Current affiliation: Development Real World Evidence, Oncology, Sanofi Inc., France

Scan here to download PDF

INTRODUCTION

- Erectile dysfunction (ED), a cardiovascular (CV) disease affecting the sexual activity is estimated to affect 322 million men by 2025, worldwide.[1,2]
- Tadalafil, a phosphodiesterase type 5 (PDE5) inhibitors, is the first-line therapy option for the treatment of ED.[3]
- PDE5 inhibitors have brought about a significant shift in the treatment of ED, however, there is a need to better understand the various demographics, comorbidities, and treatment pattern of tadalafil users suffering from ED that can aid clinicians in personalizing the treatment based on the patient need in real-world setting.[4-6]

OBJECTIVE

- To characterize the profiles of ED patients prescribed with tadalafil by clustering ED population into homogenous subpopulations with different ED presentations and risks.
- To establish a target population for education and lifestyle interventions.

METHODS

- This is a non-comparative, retrospective real-world study analysis of Optum Humedica Electronic Health Record (EHR) data (2012-2019).
- The index date (ID) for each patient was the most recent date in the patient history before January 01, 2019 when a tadalafil prescription and an ED diagnosis occurred on the same day.

Inclusion Criteria

- Men aged ≥18 years with ED diagnosis
- Prescribed tadalafil on the same day as ED diagnosis
- ≥1 year of medical history (medications prescribed and administered; immunizations; allergies; lab results [including microbiology]; habits (smoking, etc.); vital signs) prior to the ID

Exclusion Criteria

- Patients who died prior to the ID

Endpoints

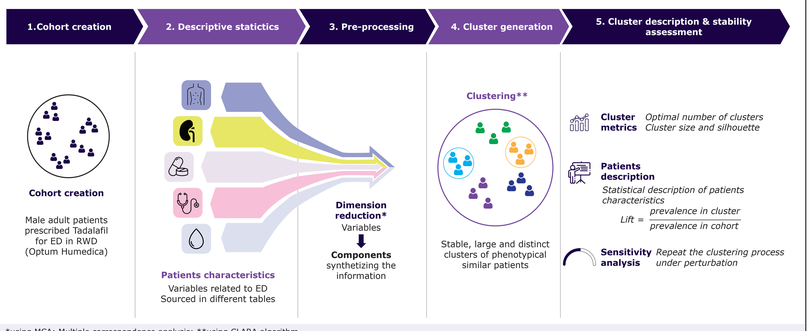
- A clustering analytical pipeline (Qluster 7) was used to segment patients into phenotypically similar clusters using literature-derived variables.[7]
- To identify and characterize the generated subpopulations from the study population based on the presence or absence of comorbidities.

Data Analyses (Figure 1)

- Optum Humedica EHR dataset used natural language processing to extract essential information from physician's notes.
- Cluster analysis using unsupervised machine learning algorithms was used to group sets of patients that were more similar with each other than to those in other groups.
- Since the ED population is likely to be highly heterogeneous, cluster analysis can identify homogenous subpopulations with different ED presentations and risks.
- Number of symptoms was calculated on the list of variables considered for the study.
- Data were analyzed in several consecutive steps following a robust analytical method:

1. **Cohort creation**
   - Prevalence of clinical variables was calculated
2. **Pre-processing literature variables**
   - Variables previously identified in the literature that were confirmed with physicians for their relevance for analysis were synthesized and refined via Multiple Correspondence Analysis.
3. **Unsupervised learning to generate patient clusters**
   - Phenotypically similar groups of patients were identified using the clustering algorithm CLARA
4. **Statistical and clinical description of clusters**
   - Lift (ratio between the prevalence within the cluster and within the cohort) of every clinical characteristic was calculated
5. **Stability assessment**
   - By repeating the clustering process under perturbations (i.e., noise and resampling) multiple times to assess results robustness

Figure 1. Overview of the data analysis steps[7]



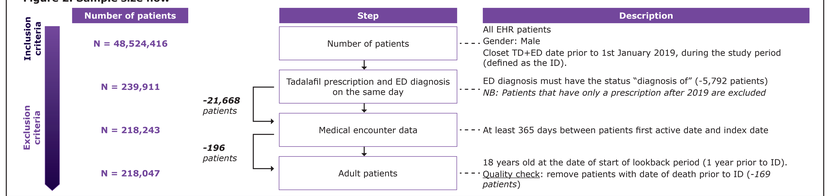
*using MCA: Multiple correspondence analysis; **using CLARA algorithm
ED, erectile dysfunction; RWD, real-world data

RESULTS

Sample Size

- Among the 48,524,416 male patients identified, 218,047 patients were included based on the fulfillment of inclusion/exclusion criteria (**Figure 2**)

Figure 2: Sample size flow



ED, erectile dysfunction; EHR, electronic health record; ID, index date; Q, quarter; TD, tadalafil.

Baseline demographic characteristics: Study population

- Baseline demographics characteristics of study population are presented in **Table 1**. Mean age was 57.1 years and 81% were Caucasian.
- Patients who had previously been prescribed tadalafil two years prior to their ID made up 42% of the cohort.
- Risk factors such as hypertension, hyperlipidemia, obesity, diabetes mellitus, and smoking were most common in these individuals.
- Prior to the ID, approximately 10% of the patients in the study population experienced a CV event.
- In addition, various conditions with a prevalence of >10% were also found, including anemia, chronic kidney disease (CKD), anxiety, depression, benign prostate hyperplasia (BPH), lumbar spine problems, and respiratory disorders.

Table 1. Baseline characteristics and demographics of the study population

Characteristic	Study population N = 218,047
Age, years, mean (SD)	57.1 (11.1)
Age group, n (%)	
18-19	40 (0.0)
20-29	2,346 (1.1)
30-39	11,554 (5.3)
40-49	38,044 (17.4)
50-59	73,506 (33.7)
60-69	64,350 (29.5)
70-79	24,156 (11.1)
80-89	4,051 (1.9)
Coverage*, years, mean (SD)	8.5 (2.8)

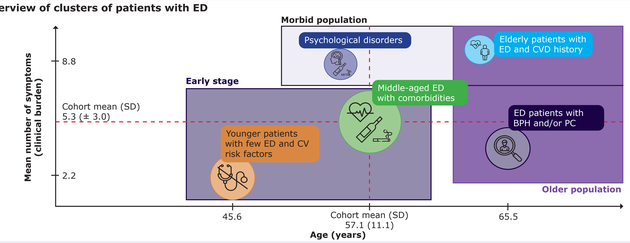
Characteristic	Study population N = 218,047
Race, n (%)	
African American	28,068 (12.9)
Asian	2,798 (1.3)
Caucasian	175,810 (80.6)
Other	11,371 (5.2)
Number of symptoms, mean (SD)	5.3 (3)
Previous prescription of tadalafil, n (%)	
2 years before ID (excluded)	91,324 (41.9)
1 year after ID (excluded)	53,247 (24.4)
Patients lost 1 year after ID, n(%)	25,810 (11)

*The number of years between first and last date active
CV, cardiovascular; ID, index date; SD, standard deviation

Clustering overview and medical characterization

- Five clusters were identified (**Figure 3**)
  - **3 clusters were homogeneous with clear distinction from each other indicating that clusters were clearly separated from each other**
    - Younger patients with few comorbidities linked to ED or other CV risk factors
    - BPH and/or prostate cancer (PC)
    - Middle-aged ED with comorbidities
  - **2 clusters were less homogenous with overlapping characteristics of other clusters**
    - Elderly patients with ED and CV history (some overlapped with the middle-aged ED with comorbidities cluster and others with those in the ED patients with BPH and/or PC cluster)
    - Psychological disorders (some overlapped with those in middle-aged ED cluster)

Figure 3: Overview of clusters of patients with ED



BPH, benign prostate hypertrophy; CV, cardiovascular; CVD, cardiovascular disease; ED, erectile dysfunction; PC, prostate cancer; SD, standard deviation.

Five distinct phenotype clusters, literature variables, prevalence and lift are summarized in **Table 2**. These clusters comprised patients with clinically distinct phenotypes ranging from different demographics and potentially disease stages (younger vs older patients) to differences in etiology (psychological disorders vs CV origin of ED)

- ***Younger patients with few ED and CV risk factors***
  - This cluster had 46,526 patients, mean age was 48.6 years, and 80.5% were Caucasian.
  - Mean number of symptoms*/patients was 2.1, 27.5% of patients had previous prescription of tadalafil (2 years before ID) and none had CV healthcare utilization.
- ***ED patients with BPH and/or PC***
  - This cluster contained 39,428 patients, mean age was 65.3 years, and 83.0% were Caucasian.
  - Mean number of symptoms*/patients was 4.4, 57.7% patients had previous prescription of tadalafil (2 years before ID), 0.2% had CV healthcare utilization.
- ***Middle-aged ED with comorbidities***
  - This cluster had 68,418 patients, mean age was 55.7 years, and 78.8% were Caucasian.
  - Mean number of symptoms*/patients was 5.0; 43.8% patients had previous prescription of tadalafil (2 years before ID), 0.3% had CV healthcare utilization.
- ***Elderly patients with ED and CV disease (CVD) history***
  - This cluster contained 29,963 patients, mean age was 65.5 years, and 76.3% were Caucasian.
  - Mean number of symptoms*/patients was 8.8; 35.7% patients had previous prescription of tadalafil (2 years before ID) 6.9% had CV healthcare utilization.
- ***Psychological disorders***
  - This cluster contained 43,712 patients, mean age was 54.5 years, and 85.5% were Caucasian.
  - Mean number of symptoms*/patients was 7.9; 44.9% patients had previous prescription of tadalafil (2 years before ID), 1.7% had CV healthcare utilization.

* The clinical burden was the number of various symptoms, up to a maximum of 36 clinical variable were selected, without considering any treatments.

Table 2: Five distinct phenotype clusters: Literature variables, prevalence and lift

		Younger patients with few ED or CV risk factors (n=46,526)		ED patients with BPH and/or PC (n=39,428)		Middle-aged ED with comorbidities (n=68,418)		Elderly patients with ED and CVD history (n=29,963)		Psychological disorders (n=43,712)	
Literature variable	**Study population prevalence (%)**	**Cluster prevalence (%)**	**Lift (% cluster/ % Study population)**	**Cluster prevalence (%)**	**Lift (% cluster/ % Study population)**	**Cluster prevalence (%)**	**Lift (% cluster/ % Study population)**	**Cluster prevalence (%)**	**Lift (% cluster/ % Study population)**	**Cluster prevalence (%)**	**Lift (% cluster/ % Study population)**
Behavioral disorders	64.7	45.4	0.7	53.0	0.8	69.0	1.1	76.8	1.2	85.7	1.3
CV factors	82.8	43.4	0.5	84.4	1.0	94.6	1.1	99.7	1.2	96.3	1.2
Demographics: Age, years											
18–44	13.0	38.0	2.9	0.3	0.0	5.8	0.4	2.5	0.2	17.2	1.3
45–54	26.3	33.5	1.3	5.8	0.2	37.6	1.4	11.1	0.4	31.1	1.2
55–64	35.3	20.3	0.6	38.4	1.1	46.8	1.3	27.4	0.8	36.3	1.0
> 65	25.4	8.2	0.3	55.5	2.2	9.8	0.4	59.1	2.3	15.4	0.6
Endocrine, nutritional, and metabolic disorders	82.4	47.1	0.6	79.7	1.0	97.9	1.2	97.7	1.2	89.4	1.1
History of CV	7.8	0.4	0.0	1.7	0.2	2.2	0.3	41.9	5.4	6.3	0.8
Kidney diseases	20.7	3.7	0.2	16.8	0.8	9.2	0.4	70.9	3.4	27.5	1.3
Neurologic and psychologic disorders	28.0	18.7	0.7	10.6	0.4	11.7	0.4	27.6	1.0	94.9	3.4
Other disorders*	58.9	25.6	0.4	82.6	1.4	48.6	0.8	78.0	1.3	81.2	1.4
Pelvic disorder	1.7	0.2	0.1	4.9	2.9	0.4	0.2	2.7	1.6	1.5	0.9

*Bacterial diseases, BPH, hyperthyroidism, low testosterone, lumbar issues, peyronies disease, prostate cancer, pulmonary hypertension, respiratory disorders, spine injury
Lift: Ratio between the prevalence within the cluster and within the study population. Literature variables: The clinical characteristics linked with the pathophysiology of ED from the literature.
BPH, benign prostate hypertrophy; CV, cardiovascular; CVD, CV disease; ED, erectile dysfunction; PC, prostate cancer.

STRENGTHS & LIMITATIONS

- The strengths of this study are that it utilizes a substantial database of real-world data, which is further improved by incorporating detailed physician's notes. The study also employs unsupervised profiling to generate new knowledge on medical expertise. Clusters sensitivity analysis was conducted to evaluate the robustness of the results and establish confidence in the subgroups within the target population. A data-driven approach combined with clinical expertise was used to select the variables.
- However, the study has some limitations, including the use of retrospective data, the possibility of selection bias, incomplete data, potential geographic bias, underrepresentation of certain ethnic groups, exclusion of other PDE5 inhibitors, and the assessment of only a momentary snapshot of ED.

CONCLUSIONS

- This novel, real-world study segmentized a heterogenous population of tadalafil-prescribed ED patients into 5 distinct phenotype clusters
  - The psychological disorders cluster suggests that certain men might benefit from extra treatments.
  - The younger patients with few comorbidities linked to ED and CV risk factors cluster indicating that they may benefit from longer treatment periods exceeding 2-5 years to reduce their risk of developing CVD.
- The data may provide guidance in curating individualized treatment options for those with various phenotypes.
- A better perspective on the association between ED, use of tadalafil, concomitant medications, and CV risk may be achieved based on these findings.
- Further studies using more data sources are warranted to strengthen the study findings.

Conflicts of Interest

KTMcV has received fees for advisory boards (Prodeon) and speaker's bureau (Antares Pharma, Inc.,), has received a grant (STTR grant) from and is a consultant/advisor for NIDDK; he is the Secretary of the Society of Benign Prostatic Disease Board of Directors; he has received honoraria from Boston Scientific; he is an Editor for UptoDate®. RAK received consultancy fees from Sanofi for studies related to switching from tadalafil to an over-the-counter product, and has received grant support from Sanofi to study the effect of PDE5is on CV outcomes in a large insurance database. **MM, MR, PG, NL,** and **CE** are employees of Quinten, which was under contract to Sanofi to conduct this work. **CD, AS, MG,** and **TMcG** are employees of Sanofi and hold shares and/or stock options in the company.

Funding

This study was funded by Sanofi.

Acknowledgement

The authors thank Amessou Mohamed for critically reviewing the poster for scientific content; Avinash Bardia, and Ashwitha A, employees of Sanofi, for writing and editorial support; and Naveen Kandibanda for poster design support.

References

1. NIH Consensus Conference. *JAMA*. 1993;270:83–90.
2. Ayta IA, et al. *BJU Int*. 1999;84:50–56.
3. Madeira CR, et al. *World J Urol*. 2021;39:953–962.
4. Katkade VB, et al. *J Multidiscip Healthc*. 2018;11:295–304.
5. Roberts MH, et al. *Pharmacoecon Open*. 2021;5:3–11.
6. Beaulieu-Jones BK, et al. *Clin Pharmacol Ther*. 2020;107:843–852.
7. Esnault C, et al. *Front Artif Intell*. 2023;5:1055294.

Poster presented at ISPOR Annual Meeting, Boston, MA, USA; May 07–10, 2023.