

Title

TreatmentPatterns: An R package to analyze treatment patterns of a study population of interest

Authors

Aniek F. Markus^{1*} MSc

Katia M.C. Verhamme^{1,2} PhD, MD

Jan A. Kors¹ PhD

Peter R. Rijnbeek¹ PhD

¹ Department of Medical Informatics, Erasmus University Medical Center, Rotterdam, The Netherlands

² Department of Bioanalysis, Ghent University, Ghent, Belgium

* Corresponding author at: Department of Medical Informatics, Erasmus University Medical Center, Rotterdam, The Netherlands. Email address: a.markus@erasmusmc.nl (A.F. Markus)

Date

December 23, 2021

Keywords

Treatment pathways, sunburst plot, current treatment practices, common data model, R package

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

ABSTRACT

Background and objectives

There is an increasing interest to use real-world data to illustrate how patients with specific medical conditions are treated in real life. Insight in the current treatment practices helps to improve and tailor patient care. We aim to provide an easy tool to support the development and analysis of treatment pathways for a wide variety of medical conditions.

Methods

We formally defined the process of constructing treatment pathways and developed an open-source R package TreatmentPatterns (<https://github.com/mi-erasmusmc/TreatmentPatterns>) to enable a reproducible and timely analysis of treatment patterns.

Results

The developed package supports the analysis of treatment patterns of specific populations of interest. We demonstrate the functionality of the package by analyzing the treatment patterns of three common chronic diseases (type II diabetes mellitus, hypertension, and depression) in the Dutch Integrated Primary Care Information (IPCI) database.

Conclusion

TreatmentPatterns is a tool to make the analysis of treatment patterns more accessible, more standardized, and more interpretation friendly. This tool can facilitate and contribute to the accumulation of knowledge on real-world treatment patterns across disease domains. We encourage researchers to further adjust and add custom analysis to the package based on their research needs.

BACKGROUND

There is an increasing interest to use real-world data to illustrate how patients with specific medical conditions are treated in real life. Real-world data is already used for post-market safety surveillance of medicines, to monitor drug consumption and related costs, to detect inappropriate prescribing and to assess compliance with treatment guidelines [1]. This research focusing on treatment patterns allows to analyze drug utilization beyond the volume of drug uptake (i.e. incidence and prevalence numbers on prescribing/dispensing), as it allows to investigate treatment pathways (i.e. first-line and subsequent treatments over time). The latter is crucial to investigate adherence to treatment guidelines in their entirety. Clinical guidelines are systematically developed recommendations informed by a review of scientific evidence and an assessment of the benefits and harms of alternative treatments for a specific medical condition [2]. These guidelines therefore support evidence-based medicine by serving as a ‘best practice’. Unfortunately, clinical guidelines are often not well adhered to in practice for a variety of reasons including lack of applicability, organizational constraints, or lack of awareness [3].

Knowledge of the current treatment practices is an important starting point to improve clinical practice and from a community perspective to facilitate the rational use of drugs. Evidence from real-world data has driven changes in clinical guidelines, but it is recognized that this data is not yet used to its full potential [1]. An increasing amount of routine health care data is available in the form of electronic health records, claims data and registries, which allows the analysis of treatment pathways in real life [4-7]. Currently, there is no easy-to-use tool available that is still flexible enough to accommodate specific research needs. Barriers to analyze treatment practices might be a lack of data interoperability and a high-level of required resources (i.e. amount of time and level of expertise).

In this paper we present TreatmentPatterns, an open-source R package to enable the analysis of treatment patterns that overcomes these barriers. The Observational Medical Outcomes Partnership (OMOP) common data model (CDM) maps data in different structures, formats and terminologies to one common standard to promote interoperability of health care data [8]. This allows to perform multi-database studies without the need to share patient-level data. Moreover, the OMOP CDM allows for the development of standardized tools, which is important to produce reproducible and timely real-world evidence. A wide range of tools is already available to support research on the OMOP CDM; Health Analytics Data-to-Evidence Suite (HADES)¹ is a set of open source R packages for large-scale analytics, including population characterization, population-level causal effect estimation, and patient-level prediction. It also provides supporting packages for new analytical pipelines. These building blocks are used by TreatmentPatterns to analyze treatment patterns of a study population of interest. The package can be executed against the OMOP CDM, but the main parts of the package are also usable with other data formats.

OBJECTIVES

We aimed to provide a user-friendly tool to support the development and analysis of treatment pathways for a wide variety of medical conditions. By making the analysis of treatment patterns more accessible, more standardized, and more interpretation friendly, we aim to contribute to the accumulation of knowledge on real-world treatment patterns across disease domains. We demonstrate the package by analyzing the treatment patterns of three common chronic diseases (type II diabetes mellitus, hypertension, and depression) using data from the Dutch Integrated Primary Care Information (IPCI) database².

¹ <https://github.com/OHDSI/Hades>

² <https://www.ipci.nl/>

METHODS

In this section, we first formally define the problem of constructing treatment pathways following earlier work [4]. Based on the problem formalization, we then describe the implementation of the open-source R package TreatmentPatterns, which enables a reproducible and timely analysis of treatment patterns.

Problem formalization

We defined the problem as follows: for a specified study population (i.e. target cohort), the goal is to find the treatment pathway consisting of selected “treatments of interest” (i.e. event cohorts).

Treatments of interest (i.e. events) can be drug prescribing and/or dispensing, but could as well be other therapies, procedures, or measurements. The treatment pathway is defined as the sequence of treatments of interest over time. The target cohort is the set of patients for whom we want to study treatment pathways and should be defined by a clear set of inclusion/exclusion criteria.

Similarly, event cohorts specify which patients have received a given treatment of interest. We use the term event era to refer to the span of time a patient is exposed to a specific treatment of interest. How to define target and event cohorts is elaborated on in other sources (e.g. Chapter 10 ‘Defining cohorts’ in the Book Of OHDSI [9]). The challenge addressed in this paper is how the treatments of interest should be processed to construct treatment pathways.

For each individual in a target cohort, we extract the treatments from his or her medical file after the index date. Note that individuals can have many different, sometimes simultaneously occurring treatments in their medical file. The index date is the moment in time the individual enters the target cohort. Figure 1 shows an example medical file for an individual receiving treatments A, B and C. The figure visualizes the decisions that need to be made to construct the treatment pathway for this individual.

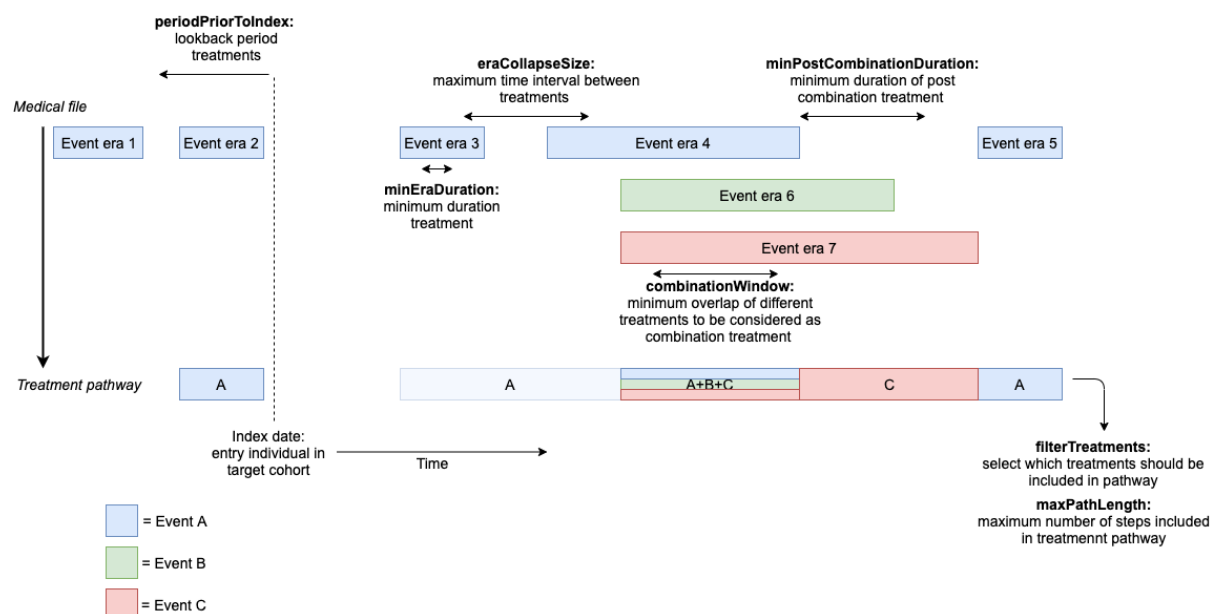


Figure 1: Summary of decisions to construct individual treatment pathways. The events A, B, and C refer to the treatments of interest. For full detailed steps from medical file to treatment pathway see Appendix A.

To create treatment pathways, the following decisions need to be made:

1. Lookback period treatments (*periodPriorToIndex*)

The first decision that needs to be made is the moment in time from which selected treatments of interest should be included in the treatment pathway. The default is all treatments starting after the index date of the target cohort. For example, for a target cohort consisting of newly diagnosed patients, treatments after the moment of first diagnosis are included. However, it might also be desirable to include (some) treatments prior to the index date, for instance in case a specific disease diagnosis is only confirmed after initiating treatment. Therefore, *periodPriorToIndex* specifies the period (i.e. number of days) prior to the index date from which treatments should be included.

2. Minimum duration treatment (*minEraDuration*)

The duration of extracted event eras may vary a lot and it can be preferable to limit to only treatments exceeding a minimum duration. Hence, *minEraDuration* specifies the minimum time an event era should last to be included in the analysis.

3. Maximum time interval between treatments (*eraCollapseSize*)

If an individual receives the same treatment for a longer period of time (e.g. need of chronic treatment), one is likely to need refills. As patients are not 100% adherent, there might be a gap between two subsequent event eras. Usually, these eras are still considered as one treatment episode and the *eraCollapseSize* defines the maximum gap within which two eras of the same event cohort would be collapsed into one era (i.e. seen as continuous treatment instead of a stop and re-initiation of the same treatment).

4. Minimum overlap of different treatments to be considered as combination treatment (*combinationWindow*)

Patients often receive different treatments at the same time, which may occur with full overlap or partial overlap. If the period of overlap is short, this is defined as a *switch* in treatment, which means that the earlier treatment is switched to a new treatment. If the overlap is longer, we assume that both treatments are received at the same time. To differentiate between the two options, *combinationWindow* specifies the time that two event eras need to overlap to be considered a combination treatment. If there are more than two overlapping event eras, we sequentially combine treatments, starting from the first two overlapping event eras.

5. Minimum duration of post combination treatment (*minPostCombinationDuration*)

As a result of creating combination treatments, there might be short remaining periods of a single treatment before or after the combination treatments. For similar reasons as described above, we

might only want to consider treatments if they last for a sufficient period of time. Therefore, *minPostCombinationDuration* defines the minimum time that an event era before or after a generated combination treatment should last to be included in the pathway as a separate treatment.

6. Select which treatments should be included in pathway (*filterTreatments*)

Depending on the research question, we might be interested to include repeated treatments in the pathway or not. The *filterTreatments* parameter allows to specify whether to include all treatments ('All'; A – A – A+B+C – C – A), to remove sequential repeated treatments ('Changes'; A – A+B+C – C – A), or to only include first time occurrences of treatments ('First'; A – A+B+C – C).

7. Maximum number of treatments included in pathway (*maxPathLength*)

Finally, *maxPathLength* specifies the maximum number of treatments included in the treatment pathway (i.e. the number of layers in a sunburst plot).

Implementation

The R package *TreatmentPatterns* implements the problem formalization described above and is structured as shown in Figure 2. The package allows to run the full analysis at once with the *executeTreatmentPatterns* function or by sequentially running the individual subfunctions (i.e. *createCohorts*, *cohortCharacterization* (optional for databases mapped to the OMOP CDM), *constructPathways*, *generateOutput*, and *launchResultsExplorer*).

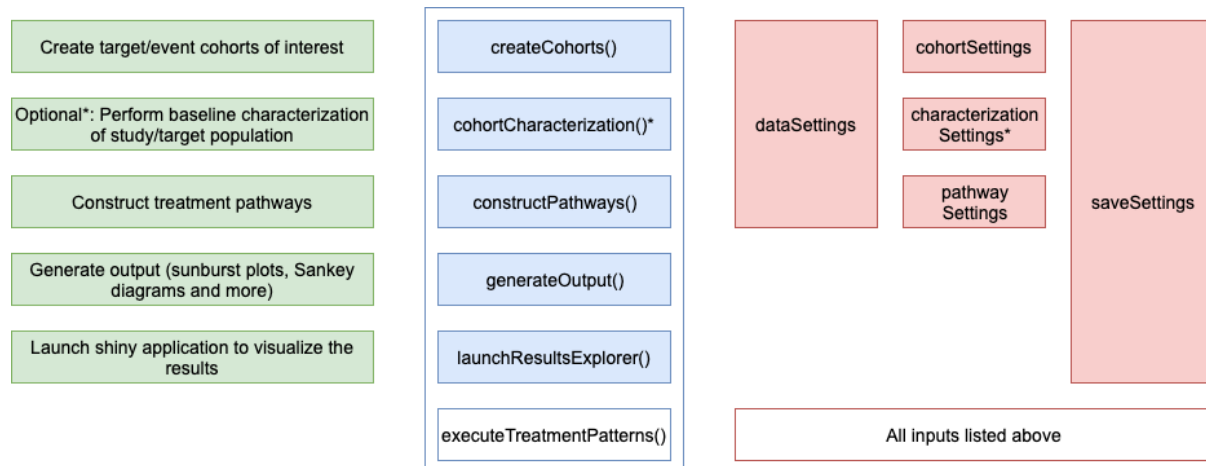


Figure 2: Structure of R package TreatmentPatterns. * Only for data in OMOP CDM format.

The main inputs of the package are the target and event cohorts of interest, which can be cohort definitions to allow automatic cohort extraction from databases mapped to the OMOP CDM or a csv file to directly import existing target and event cohorts (to be specified in *cohortSettings*). The package also allows to investigate baseline characterization of the study population including a wide range of pre-specified covariates (e.g. age/gender, Charlson comorbidity index score, comorbid conditions with different look-back periods), which can be specified in *characterizationSettings*. Moreover, it is possible to add custom covariates using SQL code. Furthermore, the *pathwaySettings* for the construction of treatment pathways need to be specified. In addition to the settings related to the creation of the individual treatment pathways above, these include output settings to aggregate the treatment pathways (e.g. *minCellCount*, *groupCombinations*). For a list of pathway settings with description and expected input we refer to Appendix B. Other necessary inputs are *dataSettings* and *saveSettings*, which indicate the location of the input data and output respectively.

After cohort extraction (and optional baseline characterization of the study population), the individual treatment pathway for each patient is constructed as follows:

1. Pre-process the medical file according to, amongst others, *periodPriorToIndex*, *minEraDuration*, and *eraCollapseSize*.

2. Identify which event eras have overlap for each person and select the first two overlapping event eras of different events per person.
3. We now distinguish three cases of overlap (see Figure 3):
 - a. If the overlap of the two event eras is smaller than *combinationWindow* AND the overlap is not equal to the total length of one of the event eras, then this is considered a *Switch*. The end date of the earlier event will be adapted to the start of the later event.

If the overlap of the two event eras is at least *combinationWindow* OR the overlap is equal to the total length of one of the events, this is considered:

- b. *FRFS* ('First Received, First Stopped') if the end date of the earlier starting event era is before or at the end date of the later starting event era.
 - c. *LRFS* ('Last Received, First Stopped') if the end date of the earlier starting event era is after the end date of the later starting event era.
4. Filter the event eras before or after a generated combination treatment based on *minPostCombinationDuration* and go back to Step 2. Stop if there are no more event eras with overlap.

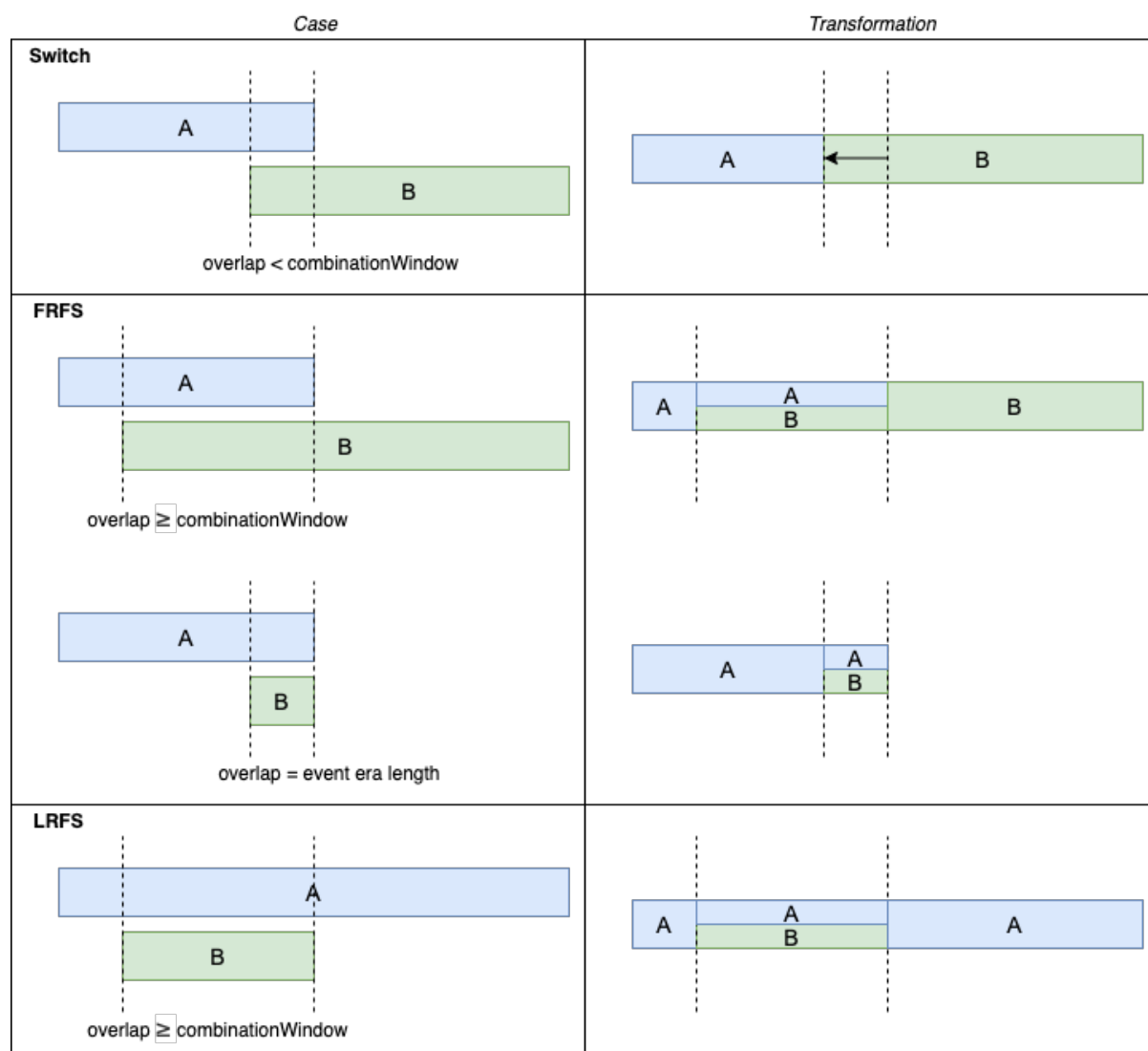


Figure 3: Three cases of overlap can occur: a) *Switch*, b) *FRFS* ('First Received, First Stopped'), and c) *LRFS* ('Last Received, First Stopped').

After constructing the individual treatment pathways, the results are aggregated for each target cohort. The package provides several functions for displaying the results of the treatment patterns analysis in both graphical and tabular formats. The package creates sunburst plots, Sankey diagrams, and various other outputs (e.g. percentage of people treated, treatment changes over time, average duration of event eras) to give insight in first-line, second-line, and higher-line treatments. *Sunburst plots* of treatment patterns show the first treatment in the center and subsequent treatments in the surrounding outer layers. Each color represents a treatment (as defined by the event cohorts) and a

layer with multiple colors indicates a combination therapy. This shows the full treatment pathway of patients over time. *Sankey diagrams* also visualize treatment patterns, but show follow-up treatment conditional on the previously received treatment by directed arrows from left to right. The size of the arrows is proportional to the number of patients. All results can be explored in an interactive Shiny application.

For full details and instructions to use the package (including a vignette and package manual) we refer to GitHub: <https://github.com/mi-erasmusmc/TreatmentPatterns>.

RESULTS

To demonstrate the functionality of the package, we present an example study analyzing treatment patterns of three common chronic diseases using data from the Dutch IPCI database. The IPCI database contains records of more than two million patients enrolled at selected general practitioners (GPs) throughout the Netherlands [10]. We aimed to study treatment patterns of patients with type II diabetes mellitus, hypertension, and depression.

Step 1: define target/event cohorts

We created three target cohorts of patients with index date at first diagnosis of the respective disease. Patients needed to have at least 365 days prior observation database time and 1,095 days follow-up time after index date (to allow sufficient time to investigate treatment patterns). We only included adult patients with a first diagnosis since 2010. The event cohorts include frequently prescribed drugs for each of the chronic diseases. An overview of the study is provided in Table 1, for more details and code lists we refer to Appendix C.

Table 1: Overview of example study on type II diabetes mellitus, hypertension, and depression.

	Type II diabetes mellitus	Hypertension	Depression
Target cohorts*	Adults with condition occurrence of: - Diabetes mellitus (201820) AND - No type 1 diabetes (35506621)	Adults with condition occurrence of: - Hypertensive disorder (316866)	Adults with condition occurrence of: - Depressive disorder (440383) AND - No schizophrenia (435783) - No bipolar disorder (432876)
Event cohorts*	Alpha-glucosidase inhibitors Biganides Dipeptidyl peptidase 4 inhibitors (DPP-4) Glucagon-like peptide-1 receptor agonists (GLP-1) Insulin Meglitinides Other anti-diabetics Sodium-glucose co-transporter-2 inhibitors (SGLT2) Sulfonylureas Thiazolidinediones	Angiotensin-converting enzyme inhibitors (ACE) Alpha-1 blockers Angiotensin receptor blockers (ARBs) Beta blockers Calcium channel blockers (CCBs) Direct vasodilators Diuretics	Norepinephrine-dopamine reuptake inhibitors (NDRIs) Selective serotonin reuptake inhibitors (SSRIs) Serotonin antagonist and reuptake inhibitors (SARIs) Serotonin-norepinephrine reuptake inhibitors (SNRIs) Tetracyclic antidepressants Tricyclic antidepressants (TCAs)
Baseline characteristics	Age, sex, Charlson Comorbidity Index score, and a history of the following conditions: cerebrovascular disease, depressive disorder, diabetes mellitus, heart failure, hypertensive disorder, ischemic heart disease, kidney disease, obesity		
Pathway settings	periodPriorToIndex = 0 days minEraDuration = 5 days eraCollapseSize = 30 days combinationWindow = 30 days minPostCombinationDuration = 30 days filterTreatments = Changes, remove sequential repeated treatments maxPathLength = 5 minCellCount = 5 minCellMethod = Adjust groupCombinations = 10 addNoPaths = TRUE		
Data source	Integrated Primary Care Information (IPCI), Netherlands		

* For more details and code lists we refer to Appendix C.

Step 2: specify baseline characteristics of interest

We chose to include the following baseline characteristics of interest: age, sex, Charlson Comorbidity Index score, and a history of the following conditions: cerebrovascular disease, depressive disorder, diabetes mellitus, heart failure, hypertensive disorder, ischemic heart disease, kidney disease, obesity. For each target cohort we excluded the history of the corresponding disease from the list of baseline characteristics.

Step 3: specify settings to construct treatment pathways

We constructed treatment pathways from the moment of first diagnosis, including the treatments of interest as specified for the different study populations in Table 1. We restricted the analysis to event eras that last at least 5 days (*minEraDuration*) and considered a treatment to be a combination if the overlap is 30 days (*combinationWindow*). We included all changes between treatments (*filterTreatments*), thereby removing sequential repeated treatments but keeping recurring treatments in the pathways (i.e. A – A+B+C – C – A). For the full list of pathway settings see Table 1.

Step 4: execute study

The complete treatment patterns study was executed by running the *executeTreatmentPatterns* function after defining the setting objects (see Appendix D for a code example). This set up has the advantage that limited programming expertise is required as the package will automatically execute all steps and produce an interactive Shiny application, the results of which are subsequently discussed.

Step 5: check out results

The results of the example study are available at:

<https://aniekmarkus.shinyapps.io/TreatmentPatterns/>. We found 50,285 type II diabetes mellitus patients, 120,675 hypertension patients, and 32,567 patients with depression. Baseline

characteristics showed that the proportion of males is slightly lower than that of females in patients newly diagnosed with hypertension (44.2%) and depression (35.8%) whereas for type II diabetes mellitus no female predominance was observed. The interface allows to easily compare baseline characteristics across databases or study populations (as shown in Figure 4).

Table with selected demographics and patient characteristics (in percentages).
 Show 20 entries Search:

	Covariate name	Depression	Hypertension	T2DM
1	Number of persons	32567	120675	50285
2	Male	35.8	44.2	50.3
3	Age	48.8	63.6	63
4	Charlson comorbidity index score	0.8	1	2
5	Cerebrovascular disease	4	7	7.4
6	Depressive disorder		9.1	10.5
7	Diabetes mellitus	10.2	16.2	
8	Heart failure	1.5	2.7	4.2
9	Hypertensive disorder	21.9		53.5
10	Ischemic heart disease	6	11.4	16
11	Kidney disease	4.2	6.6	8.4
12	Obesity	6.1	6.1	10.1

Showing 1 to 12 of 12 entries Previous 1 Next

Figure 4: Baseline characterization of study populations in IPCI.

Treatment pathways were visualized in the form of various plots and an example of a sunburst plot is shown in Figure 5 (for remaining plots see Appendix E). For type II diabetes mellitus, biguanides were the most frequently prescribed first-line treatments (62.4% of treated patients) followed by insulin (9.9% of treated patients). GPs thus seems to align with regard to the type of first-line treatment for type II diabetes mellitus. The proportion of switching was low as 35.2% of treated patients remained on biguanides (i.e. did not receive another type of treatment afterwards). More variability in the type of second-line treatments was observed, and this also included more combination treatments (62.3% as opposed to 11.1%). For hypertension, ACE inhibitors (26.7% of treated patients), beta blockers (19.1% of treated patients), CCBs (17.1% of treated patients) and

ARBs (16.5% of treated patients) were all frequently prescribed as first treatment (see Figure 1 in Appendix E). For depression, SSRIs were the most frequently prescribed first-line treatment (61.2% of treated patients), followed by three types of medication that were each received by at least 10% of the treated patients: SNRIs, TCAs, and tetracyclic antidepressants (see Figure 2 Appendix E). Interestingly, patients treated for depression often did not proceed to a second-line treatment (only 24.8%). Use of combination treatments as first-line treatment was mainly observed in patients with hypertension (15.0%) and type II diabetes mellitus (11.1%), but was rare in patients with depression (1.9%).

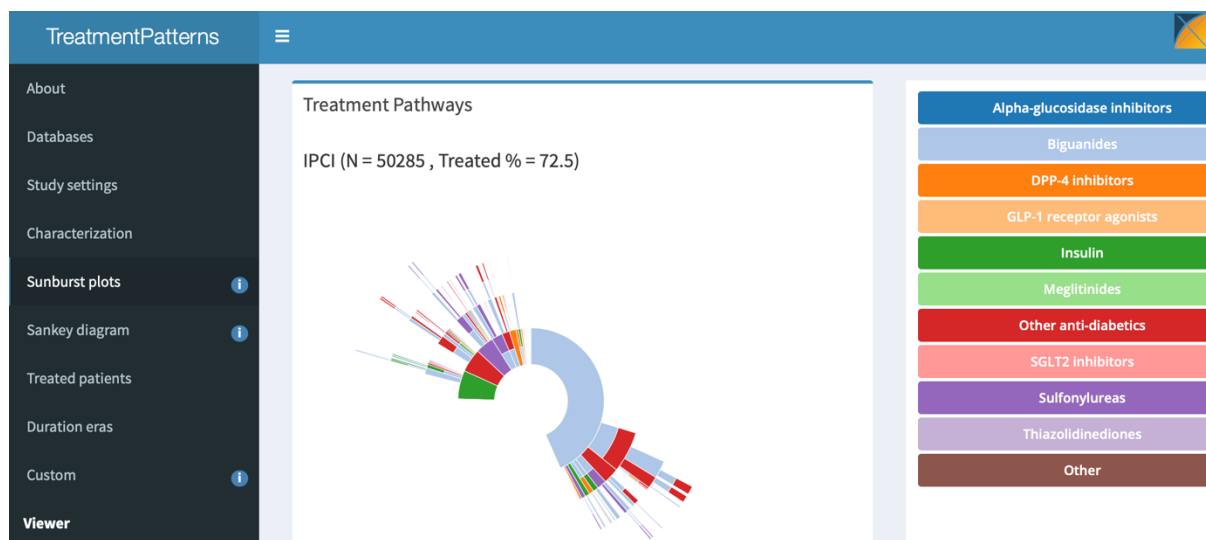


Figure 5: Sunburst plot visualizing the treatment pathways of patients with type II diabetes mellitus in IPCI.

Finally, the heatmaps indicating the duration of event eras (see Figures 6-8 in Appendix E) showed that the mean duration of first-line treatment for type II diabetes mellitus was 661 days, for hypertension 674 days, and for depression 658 days.

DISCUSSION

In this paper, we formalize the construction of treatment pathways and present an R package implementing the presented analysis. This allows researchers to give insight in treatment patterns of

a specified study population (target cohort) and selected treatments of interest (event cohorts) in a reproducible and timely manner. TreatmentPatterns facilitates the execution of these studies at a large-scale, i.e. across data sources, countries, and disease domains, while allowing to easily adjust settings and perform sensitivity analysis to investigate the robustness of the results (both within and between data sources).

Insight in the current treatment practices helps to improve clinical practice. It gives insight in the choice of treatment and preferences of doctors taking into account patient's phenotype characteristics. In addition, it provides indirect information on pharmaceutical expenditure and potential health care budget constraints. Furthermore, this provides knowledge on the extent to which clinical guidelines are adhered to, which is an important step in further implementing and/or revisiting current guidelines.

Previous work analyzing treatment pathways is often addressing specific use cases and includes for example the assessment of current treatment practices of patients with type II diabetes mellitus [4 7], hypertension [4], depression [4 5], and epilepsy [6]. Some methodological work on treatment patterns exists, but focuses for example on extracting treatment patterns using a multi-view similarity fusion method [10] or alternative visualization techniques [11]. To our knowledge, no earlier work formally defined the process of constructing treatment pathways from medical files. The already existing 'Cohort Pathways' tool³, which was used for some of the before-mentioned studies, has the disadvantage that it has a very limited number of settings and is not customizable to address specific research needs. The presented R package is thus an addition to the existing base of methods available to generate evidence from routinely collected data.

³ Available in ATLAS: <https://atlas-demo.ohdsi.org>.

TreatmentPatterns is a user-friendly tool to support the development and analysis of treatment pathways without requiring much programming. Therefore, the target audience is large and includes - but is not limited to - researchers, clinicians, the pharmaceutical industry, regulatory bodies, and health technology assessment agencies. The quality of the analysis is dependent on the target and event cohort definitions (i.e. phenotypes), which need to be created and validated by the user (e.g. using the OHDSI CohortDiagnostics tool). However, this is a limitation related to the quality of the data being used that holds for all types of observational data studies and is not specific for research investigating treatment patterns. Finally, it should be noted that the analysis of treatment patterns in observational data is limited to prescribing and/or dispensing data. Hence, it is not possible to infer actual treatment intake. This should be taken into account when interpreting the results.

In conclusion, this tool is intended to make the analysis of treatment patterns more accessible, more standardized, and more interpretation friendly. We hope it thereby contributes to the accumulation of knowledge on real-world treatment patterns across disease domains. Note that there might be disease-specific features of interest, that are not included in the current package. We thus encourage researchers to further adjust and add custom analysis to the package based on their research needs.

ACKNOWLEDGEMENTS

The authors would like to thank the Observational Health Data Sciences and Informatics (OHDSI) community for building the foundation necessary to perform large-scale, international, multi-database studies through the development of the OMOP CDM and open-source tools. In particular, we like to thank the authors of the earlier paper ‘Characterizing treatment pathways at scale using the OHDSI network’ [4], which has provided the initial foundation and rationale to develop the TreatmentPatterns package.

CONFLICTS OF INTERESTS

The authors declare that they have no competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

FUNDING

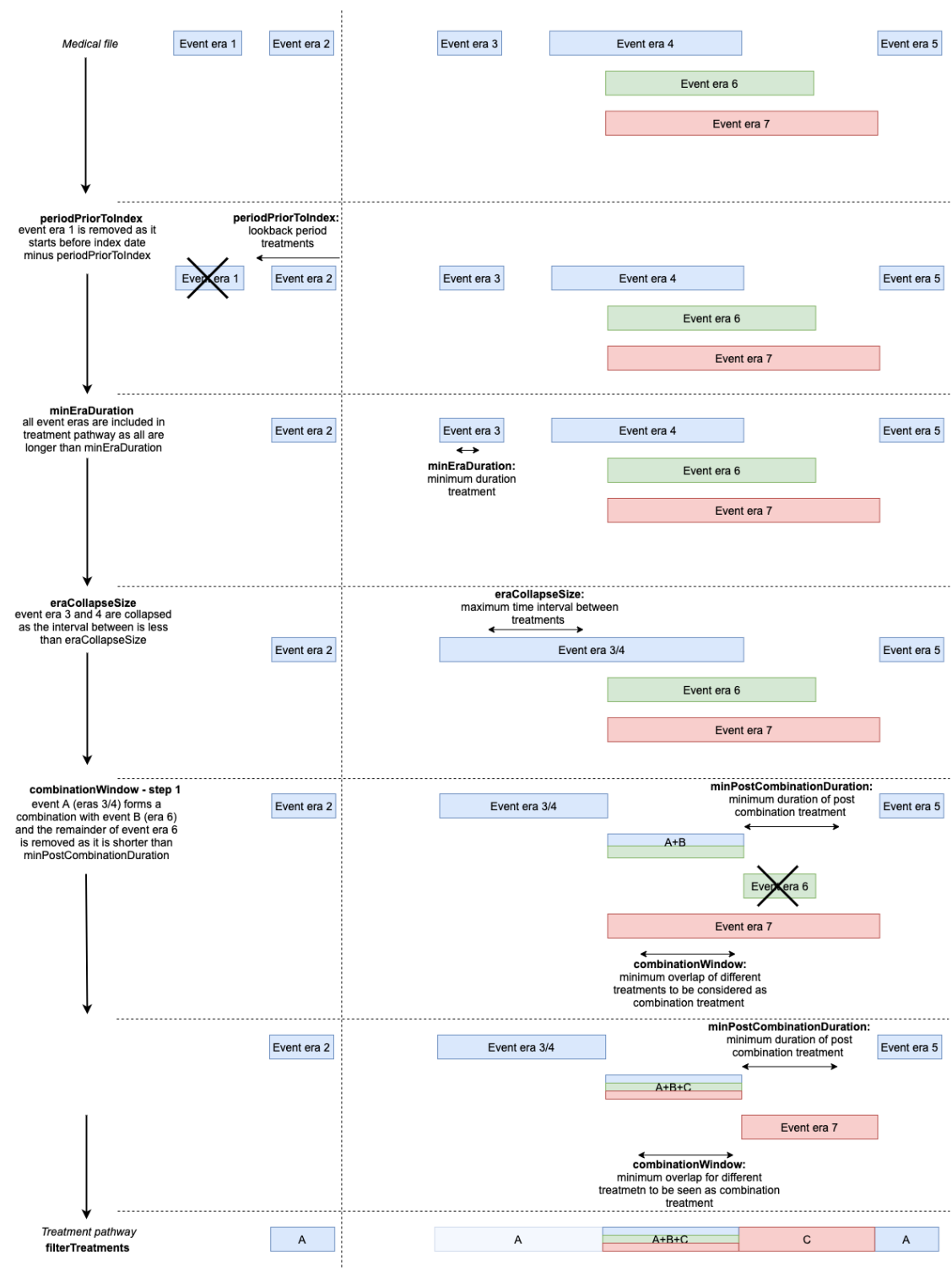
This project has received support from the European Health Data and Evidence Network (EHDEN) project. EHDEN received funding from the Innovative Medicines Initiative 2 Joint Undertaking (JU) under grant agreement No 806968. The JU receives support from the European Union's Horizon 2020 research and innovation programme and EFPIA.

REFERENCES

1. OECD. Using Routinely Collected Data to Inform Pharmaceutical Policies. 2019. <https://www.oecd.org/health/health-systems/Using-Routinely-Collected-Data-to-Inform-Pharmaceutical-Policies-Analytical-Report-2019.pdf>.
2. American Academy of Family Physicians. Clinical Practice Guideline Manual. 2017. <https://www.aafp.org/family-physician/patient-care/clinical-recommendations/cpg-manual.html#xiii>.
3. Lugtenberg M, Zegers-van Schaick JM, Westert GP, Burgers JS. Why don't physicians adhere to guideline recommendations in practice? An analysis of barriers among Dutch general practitioners. *Implementation Science* 2009;**4**(1):54 doi: 10.1186/1748-5908-4-54.
4. Hripcsak G, Ryan PB, Duke JD, et al. Characterizing treatment pathways at scale using the OHDSI network. *Proceedings of the National Academy of Sciences* 2016;**113**(27):7329-36 doi: 10.1073/pnas.1510502113.
5. Kern DM, Cepeda MS, Defalco F, Etropolski M. Treatment patterns and sequences of pharmacotherapy for patients diagnosed with depression in the United States: 2014 through 2019. *BMC Psychiatry* 2020;**20**(1):4 doi: 10.1186/s12888-019-2418-7.
6. Kim H, Yoo S, Jeon Y, et al. Characterization of Anti-seizure Medication Treatment Pathways in Pediatric Epilepsy Using the Electronic Health Record-Based Common Data Model. *Front. Neurol.* 2020;**11**(409) doi: 10.3389/fneur.2020.00409.
7. Lee KA, Jin HY, Kim YJ, Im Y-J, Kim E-Y, Park TS. Treatment Patterns of Type 2 Diabetes Assessed Using a Common Data Model Based on Electronic Health Records of 2000–2019. *J. Korean Med. Sci.* 2021;**36**(36):e230-e30
8. Overhage JM, Ryan PB, Reich CG, Hartzema AG, Stang PE. Validation of a common data model for active safety surveillance research. *J. Am. Med. Inform. Assoc.* 2012;**19**(1):54-60 doi: 10.1136/amiajnl-2011-000376.
9. Observational Health Data Sciences and Informatics. The Book of OHDSI. 2019. <https://ohdsi.github.io/TheBookOfOhdsi/>.
10. Vlug A, Van der Lei J, Mosseveld BT, et al. Postmarketing surveillance based on electronic patient records: the IPCI project. *Methods Inf. Med.* 1999;**38**(04/05):339-44 doi: 10.1055/s-0038-1634402.

SUPPLEMENTARY MATERIALS

Appendix A: Example to construct individual treatment pathway



Appendix B: Pathway settings

Table 1: List of pathway settings with description and expected input.

General settings		
studyName	Unique name identifying the set of study parameters below	Name
targetCohortId	Select one study population	Cohort ID
eventCohortIds	Select all treatments of interest	List of cohort IDs
Individual pathway settings		
periodPriorToIndex	The period (number of days) prior to the index date of the target cohort from which treatments should be included	Number of days
minEraDuration	Minimum time an event era should last to be included in the analysis	Number of days
eraCollapseSize	Maximum gap within two eras of the same event cohort which would still allow the eras to be collapsed into one era	Number of days
combinationWindow	Time that two event eras need to overlap to be considered a combination treatment	Number of days
minPostCombinationDuration	Minimum time that an event era before or after a generated combination treatment should last to be included in the pathway as a separate treatment	Number of days
filterTreatments	Select which treatments should be included in pathway first time occurrences of treatments ('First'), remove sequential repeated treatments ('Changes'), all treatments ('All')	First / Changes / All
maxPathLength	Maximum number of treatments included in pathway	Number of steps (max. 5)

Aggregate pathway settings		
minCellCount	Minimum number of persons with a specific treatment pathway for the pathway to be included in analysis	Number of subject
minCellMethod	Select to completely remove / sequentially adjust (by removing last step as often as necessary) treatment pathways below minCellCount	Remove / Adjust
groupCombinations	Select to group all non-fixed combinations in one category 'other' in the sunburst plot	TRUE / # below which to group
addNoPaths	Select to include untreated persons without treatment pathway in the sunburst plot	TRUE / FALSE

Appendix C: Example study type II diabetes mellitus, hypertension, and depression

Target cohorts:

- i) T2DM12mo

Initial Event Cohort:

People with continuous observation of 365 days before and 1,095 days after event may enter the cohort when observing any of the following:

- Condition occurrence of Diabetes* for the first time in the person's history.

* Concept ID	Concept Name	Domain Vocabulary	Excluded	Descendants	Mapped
201820	Diabetes mellitus	Condition	SNOMED	NO YES	NO

Inclusion Criteria:

- Age >= 18 years.
- No condition occurrences of Diabetes type 1** in their history or during follow-up time.

** Concept ID	Concept Name	Domain Vocabulary	Excluded	Descendants	Mapped
35506621	Type 1 diabetes mellitus	Condition	MedDRA	NO	YES NO

- Index date starting after December 31, 2009.

Cohort Exit: the person exits the cohort at the end of continuous observation.

ii) HTN12mo

Initial Event Cohort:

People with continuous observation of 365 days before and 1,095 days after event may enter the cohort when observing any of the following:

- Condition occurrence of Hypertension* for the first time in the person's history.

*Concept ID	Concept Name	Domain Vocabulary	Excluded	Descendants	Mapped
316866	Hypertensive disorder	Condition	SNOMED	NO	YES NO

Inclusion Criteria:

- Age >= 18 years.
- Index date starting after December 31, 2009.

Cohort Exit: the person exits the cohort at the end of continuous observation.

iii) Depression12mo

Initial Event Cohort:

People with continuous observation of 365 days before and 1,095 days after event may enter the cohort when observing any of the following:

- Condition occurrence of Depression* for the first time in the person's history.

*Concept ID	Concept Name	Domain Vocabulary	Excluded	Descendants	Mapped
440383	Depressive disorder	Condition	SNOMED	NO	YES NO

Inclusion Criteria:

- Age >= 18 years.

- No condition occurrences of Schizophrenia/Bipolar** in their history or during follow-up time.

**Concept ID	Concept Name	Domain Vocabulary	Excluded	Descendants	Mapped
432876	Bipolar I disorder	Condition	SNOMED	NO	YES NO
435783	Schizophrenia	Condition	SNOMED	NO	YES NO

- Index date starting after December 31, 2009.

Cohort Exit: the person exits the cohort at the end of continuous observation.

Event cohorts:

Initial Event Cohort:

People with continuous observation of 0 days before and 0 days after event may enter the cohort

when observing any of the following:

- Drug exposure of treatment of interest*.

Cohort Exit: If the index event is found within an era, the cohort end date will use the era's end date.

Otherwise, it will use the observation period end date that contains the index event.

* Treatments of interest are defined by the following concept sets, details of Concept IDs can be found on <https://athena.ohdsi.org/>:

Study	Treatment of interest	Concept IDs
Type II diabetes mellitus	Alpha-glucosidase inhibitors	1510202, 1529331
	Biguanides	1593986,1503297,19033909,40798673
	DPP-4 inhibitors	1580747,19122137,40166035,40239216,43013884
	GLP-1 receptor agonists	793143,1583722,40170911,44506754,44816332,45774435
	Insulin	1596977,1550023,1567198,1502905,1513876,1531601,1586346,1544838,1516976,1590165,1513849,1562586,1588986,1513843,1586369,35605670,35602717,21600713,19078608
	Meglitinides	1502826,1516766

	Other anti-diabetics	1529331,1530014,730548,19033498,19001409,19059796,19001441,1510202,1502826,1525215,1516766,1547504,1515249
	SGLT2 inhibitors	793293,43526465,44785829,45774751
	Sulfonylureas	1502809,1502855,1559684,1560171,1594973,1597756,19097821
	Thiazolidinediones	1547504,1525215,1515249
Hypertension	ACE inhibitors	1335471,1340128,1341927,1363749,1308216,1310756,1373225,1331235,1334456,1342439
	Alpha-1 blockers	1363053,1350489,1341238
	Angiotensin receptor blockers (ARBs)	40235485,1351557,1346686,1347384,1367500,40226742,1317640,1308842
	Beta blockers	1319998,1314002,1322081,1338005,1346823,1386957,1307046,1313200,1314577,1327978,1345858,1353766
	Calcium channel blockers (CCBs)	1332418,1328165,1353776,1326012,1318137,1318853,1319880,1307863
	Direct vasodilators	1373928,131
	Diuretics	991382,932745,1309799,956874,970250,942350,904542
Depression	Electroconvulsive therapy	2007727,2007728,2108578,2108579,2213552,4004830,4020981,4030840,4111663,4210144,4210145,4332436,4336318,44508134
	Norepinephrine-dopamine reuptake inhibitors (NDRIs)	750982
	Psychotherapy	2007730,2007731,2007746,2007747,2007748,2007749,2007750,2007763,2108571,2213520,2213521,2213522,2213523,2213524,2213525,2213526,2213527,2213528,2213529,2213530,2213531,2213532,2213533,2213534,2213535,2213536,2213537,2213538,2213539,2213540,2213541,2213542,2213543,2213544,2213546,2213547,2213548,2213549,2213550,2213554,2213555,2617477,2617478,4012488,4028920,4035812,4048385,4048387,4079500,4079608,4079938,4079939,4080044,4080048,4083129,4083130,4083131,4083133,4083706,4084195,4084201,4084202,4088889,4100341,4103512,4114491,4117915,4118797,4118798,4118800,4118801,4119334,4119335,4121662,4126653,4128268,4128406,4132436,4136352,4137086,4143316,4148398,4148765,4151904,4164790,4173581,4179241,4196062,4199042,4202234,4208314,4219683,4221997,4225

	728,4226275,4226276,4233181,4234402,4234476,4242119,4249602,4258834,4262582,4263758,4265313,4268909,4272803,4278094,4295027,4296166,4299728,4311943,4327941,40482841,43527904,43527905,43527986,43527987,43527988,43527989,43527990,43527991,44791916,44792695,44808259,44808677,45763911,45765516,45887728,45887951,45888237,45889353,46286330,46286403
Selective serotonin reuptake inhibitors (SSRIs)	797617,715939,755695,722031,739138,40234834
Serotonin antagonist and reuptake inhibitors (SARIs)	703547
Serotonin-norepinephrine reuptake inhibitors (SNRIs)	717607,715259,743670
Tetracyclic antidepressants	725131
Tricyclic antidepressants (TCAs)	710062,738156,721724

Baseline characterization:

Covariate Name	Covariate ID (= Concept ID + time window 209)
Age	1002
Charlson comorbidity index score	1901
Male	8507001
Kidney disease	198124209
Diabetes mellitus	201820209
Heart failure	316139209
Hypertensive disorder	316866209
Cerebrovascular disease	381591209
Obesity	433736209
Depressive disorder	440383209
Ischemic heart disease	4185932209

Appendix D: Code to run example treatment patterns study

```
library(TreatmentPatterns)
dataSettings <- createDataSettings(QMOP_CDM = TRUE,
  connectionDetails = DatabaseConnector::createConnectionDetails(dbms = Sys.getenv('dbms'),
    server = Sys.getenv('server'),
    user = Sys.getenv('user'),
    password = Sys.getenv('password'),
    port = Sys.getenv('port')),
  cdmDatabaseSchema = 'cdm',
  cohortDatabaseSchema = 'results',
  cohortTable = "treatmentpatterns_examplestudy")

cohortSettings <- createCohortSettings(targetCohorts = data.frame(cohortId = c(1, 2, 3),
  atlasId = c(589, 590, 591),
  cohortName = c('T2DM', 'Hypertension', 'Depression'),
  conceptSet = c("")),
  eventCohorts = data.frame(cohortId = c(10, 11, 12, 13, 14, 15, 16, 17, 18, 19,
  110, 111, 112, 113, 114, 115, 116,
  210, 211, 212, 213, 214, 215, 216, 217),
  atlasId = c(""),
  cohortName = c('Alpha-glucosidase inhibitors', 'Biguanides', 'DPP-4 inhibitors', 'GLP-1 receptor agonists',
  'ACE inhibitors', 'Alpha-1 blockers', 'Angiotensin receptor blockers (ARBs)', 'Beta-blockers',
  'Electroconvulsive therapy', 'Norepinephrine-dopamine reuptake inhibitors (NDRIs)',
  'Electroconvulsive therapy', 'Norepinephrine-dopamine reuptake inhibitors (NDRIs)'),
  conceptSet = c("1510202;1529331", "1593986;1503297;19033909;40798673", "1580747;19122137;40166035;40166036;40166037;40166038;40166039;40166040;40166041;40166042;40166043;40166044;40166045;40166046;40166047;40166048;40166049;40166050;40166051;40166052;40166053;40166054;40166055;40166056;40166057;40166058;40166059;40166060;40166061;40166062;40166063;40166064;40166065;40166066;40166067;40166068;40166069;40166070;40166071;40166072;40166073;40166074;40166075;40166076;40166077;40166078;40166079;40166080;40166081;40166082;40166083;40166084;40166085;40166086;40166087;40166088;40166089;40166090;40166091;40166092;40166093;40166094;40166095;40166096;40166097;40166098;40166099;40166100;40166101;40166102;40166103;40166104;40166105;40166106;40166107;40166108;40166109;40166110;40166111;40166112;40166113;40166114;40166115;40166116;40166117;40166118;40166119;40166120;40166121;40166122;40166123;40166124;40166125;40166126;40166127;40166128;40166129;40166130;40166131;40166132;40166133;40166134;40166135;40166136;40166137;40166138;40166139;40166140;40166141;40166142;40166143;40166144;40166145;40166146;40166147;40166148;40166149;40166150;40166151;40166152;40166153;40166154;40166155;40166156;40166157;40166158;40166159;40166160;40166161;40166162;40166163;40166164;40166165;40166166;40166167;40166168;40166169;40166170;40166171;40166172;40166173;40166174;40166175;40166176;40166177;40166178;40166179;40166180;40166181;40166182;40166183;40166184;40166185;40166186;40166187;40166188;40166189;40166190;40166191;40166192;40166193;40166194;40166195;40166196;40166197;40166198;40166199;40166200;40166201;40166202;40166203;40166204;40166205;40166206;40166207;40166208;40166209;40166210;40166211;40166212;40166213;40166214;40166215;40166216;40166217;40166218;40166219;40166220;40166221;40166222;40166223;40166224;40166225;40166226;40166227;40166228;40166229;40166230;40166231;40166232;40166233;40166234;40166235;40166236;40166237;40166238;40166239;40166240;40166241;40166242;40166243;40166244;40166245;40166246;40166247;40166248;40166249;40166250;40166251;40166252;40166253;40166254;40166255;40166256;40166257;40166258;40166259;40166260;40166261;40166262;40166263;40166264;40166265;40166266;40166267;40166268;40166269;40166270;40166271;40166272;40166273;40166274;40166275;40166276;40166277;40166278;40166279;40166280;40166281;40166282;40166283;40166284;40166285;40166286;40166287;40166288;40166289;40166290;40166291;40166292;40166293;40166294;40166295;40166296;40166297;40166298;40166299;40166300;40166301;40166302;40166303;40166304;40166305;40166306;40166307;40166308;40166309;40166310;40166311;40166312;40166313;40166314;40166315;40166316;40166317;40166318;40166319;40166320;40166321;40166322;40166323;40166324;40166325;40166326;40166327;40166328;40166329;40166330;40166331;40166332;40166333;40166334;40166335;40166336;40166337;40166338;40166339;40166340;40166341;40166342;40166343;40166344;40166345;40166346;40166347;40166348;40166349;40166350;40166351;40166352;40166353;40166354;40166355;40166356;40166357;40166358;40166359;40166360;40166361;40166362;40166363;40166364;40166365;40166366;40166367;40166368;40166369;40166370;40166371;40166372;40166373;40166374;40166375;40166376;40166377;40166378;40166379;40166380;40166381;40166382;40166383;40166384;40166385;40166386;40166387;40166388;40166389;40166390;40166391;40166392;40166393;40166394;40166395;40166396;40166397;40166398;40166399;40166400;40166401;40166402;40166403;40166404;40166405;40166406;40166407;40166408;40166409;40166410;40166411;40166412;40166413;40166414;40166415;40166416;40166417;40166418;40166419;40166420;40166421;40166422;40166423;40166424;40166425;40166426;40166427;40166428;40166429;40166430;40166431;40166432;40166433;40166434;40166435;40166436;40166437;40166438;40166439;40166440;40166441;40166442;40166443;40166444;40166445;40166446;40166447;40166448;40166449;40166450;40166451;40166452;40166453;40166454;40166455;40166456;40166457;40166458;40166459;40166460;40166461;40166462;40166463;40166464;40166465;40166466;40166467;40166468;40166469;40166470;40166471;40166472;40166473;40166474;40166475;40166476;40166477;40166478;40166479;40166480;40166481;40166482;40166483;40166484;40166485;40166486;40166487;40166488;40166489;40166490;40166491;40166492;40166493;40166494;40166495;40166496;40166497;40166498;40166499;40166500;40166501;40166502;40166503;40166504;40166505;40166506;40166507;40166508;40166509;40166510;40166511;40166512;40166513;40166514;40166515;40166516;40166517;40166518;40166519;40166520;40166521;40166522;40166523;40166524;40166525;40166526;40166527;40166528;40166529;40166530;40166531;40166532;40166533;40166534;40166535;40166536;40166537;40166538;40166539;40166540;40166541;40166542;40166543;40166544;40166545;40166546;40166547;40166548;40166549;40166550;40166551;40166552;40166553;40166554;40166555;40166556;40166557;40166558;40166559;40166560;40166561;40166562;40166563;40166564;40166565;40166566;40166567;40166568;40166569;40166570;40166571;40166572;40166573;40166574;40166575;40166576;40166577;40166578;40166579;40166580;40166581;40166582;40166583;40166584;40166585;40166586;40166587;40166588;40166589;40166590;40166591;40166592;40166593;40166594;40166595;40166596;40166597;40166598;40166599;40166600;40166601;40166602;40166603;40166604;40166605;40166606;40166607;40166608;40166609;40166610;40166611;40166612;40166613;40166614;40166615;40166616;40166617;40166618;40166619;40166620;40166621;40166622;40166623;40166624;40166625;40166626;40166627;40166628;40166629;40166630;40166631;40166632;40166633;40166634;40166635;40166636;40166637;40166638;40166639;40166640;40166641;40166642;40166643;40166644;40166645;40166646;40166647;40166648;40166649;40166650;40166651;40166652;40166653;40166654;40166655;40166656;40166657;40166658;40166659;40166660;40166661;40166662;40166663;40166664;40166665;40166666;40166667;40166668;40166669;40166670;40166671;40166672;40166673;40166674;40166675;40166676;40166677;40166678;40166679;40166680;40166681;40166682;40166683;40166684;40166685;40166686;40166687;40166688;40166689;40166690;40166691;40166692;40166693;40166694;40166695;40166696;40166697;40166698;40166699;40166700;40166701;40166702;40166703;40166704;40166705;40166706;40166707;40166708;40166709;40166710;40166711;40166712;40166713;40166714;40166715;40166716;40166717;40166718;40166719;40166720;40166721;40166722;40166723;40166724;40166725;40166726;40166727;40166728;40166729;40166730;40166731;40166732;40166733;40166734;40166735;40166736;40166737;40166738;40166739;40166740;40166741;40166742;40166743;40166744;40166745;40166746;40166747;40166748;40166749;40166750;40166751;40166752;40166753;40166754;40166755;40166756;40166757;40166758;40166759;40166760;40166761;40166762;40166763;40166764;40166765;40166766;40166767;40166768;40166769;40166770;40166771;40166772;40166773;40166774;40166775;40166776;40166777;40166778;40166779;40166780;40166781;40166782;40166783;40166784;40166785;40166786;40166787;40166788;40166789;40166790;40166791;40166792;40166793;40166794;40166795;40166796;40166797;40166798;40166799;40166800;40166801;40166802;40166803;40166804;40166805;40166806;40166807;40166808;40166809;40166810;40166811;40166812;40166813;40166814;40166815;40166816;40166817;40166818;40166819;40166820;40166821;40166822;40166823;40166824;40166825;40166826;40166827;40166828;40166829;40166830;40166831;40166832;40166833;40166834;40166835;40166836;40166837;40166838;40166839;40166840;40166841;40166842;40166843;40166844;40166845;40166846;40166847;40166848;40166849;40166850;40166851;40166852;40166853;40166854;40166855;40166856;40166857;40166858;40166859;40166860;40166861;40166862;40166863;40166864;40166865;40166866;40166867;40166868;40166869;40166870;40166871;40166872;40166873;40166874;40166875;40166876;40166877;40166878;40166879;40166880;40166881;40166882;40166883;40166884;40166885;40166886;40166887;40166888;40166889;40166890;40166891;40166892;40166893;40166894;40166895;40166896;40166897;40166898;40166899;40166900;40166901;40166902;40166903;40166904;40166905;40166906;40166907;40166908;40166909;40166910;40166911;40166912;40166913;40166914;40166915;40166916;40166917;40166918;40166919;40166920;40166921;40166922;40166923;40166924;40166925;40166926;40166927;40166928;40166929;40166930;40166931;40166932;40166933;40166934;40166935;40166936;40166937;40166938;40166939;40166940;40166941;40166942;40166943;40166944;40166945;40166946;40166947;40166948;40166949;40166950;40166951;40166952;40166953;40166954;40166955;40166956;40166957;40166958;40166959;40166960;40166961;40166962;40166963;40166964;40166965;40166966;40166967;40166968;40166969;40166970;40166971;40166972;40166973;40166974;40166975;40166976;40166977;40166978;40166979;40166980;40166981;40166982;40166983;40166984;40166985;40166986;40166987;40166988;40166989;40166990;40166991;40166992;40166993;40166994;40166995;40166996;40166997;40166998;40166999;40167000;40167001;40167002;40167003;40167004;40167005;40167006;40167007;40167008;40167009;40167010;40167011;40167012;40167013;40167014;40167015;40167016;40167017;40167018;40167019;40167020;40167021;40167022;40167023;40167024;40167025;40167026;40167027;40167028;40167029;40167030;40167031;40167032;40167033;40167034;40167035;40167036;40167037;40167038;40167039;40167040;40167041;40167042;40167043;40167044;40167045;40167046;40167047;40167048;40167049;40167050;40167051;40167052;40167053;40167054;40167055;40167056;40167057;40167058;40167059;40167060;40167061;40167062;40167063;40167064;40167065;40167066;40167067;40167068;40167069;40167070;40167071;40167072;40167073;40167074;40167075;40167076;40167077;40167078;40167079;40167080;40167081;40167082;40167083;40167084;40167085;40167086;40167087;40167088;40167089;40167090;40167091;40167092;40167093;40167094;40167095;40167096;40167097;40167098;40167099;40167100;40167101;40167102;40167103;40167104;40167105;40167106;40167107;40167108;40167109;40167110;40167111;40167112;40167113;40167114;40167115;40167116;40167117;40167118;40167119;40167120;40167121;40167122;40167123;40167124;40167125;40167126;40167127;40167128;40167129;40167130;40167131;40167132;40167133;40167134;40167135;40167136;40167137;40167138;40167139;40167140;40167141;40167142;40167143;40167144;40167145;40167146;40167147;40167148;40167149;40167150;40167151;40167152;40167153;40167154;40167155;40167156;40167157;40167158;40167159;40167160;40167161;40167162;40167163;40167164;40167165;40167166;40167167;40167168;40167169;40167170;40167171;40167172;40167173;40167174;40167175;40167176;40167177;40167178;40167179;40167180;40167181;40167182;40167183;40167184;40167185;40167186;40167187;40167188;40167189;40167190;40167191;40167192;40167193;40167194;40167195;40167196;40167197;40167198;40167199;40167200;40167201;40167202;40167203;40167204;40167205;40167206;40167207;40167208;40167209;40167210;40167211;40167212;40167213;40167214;40167215;40167216;40167217;40167218;40167219;40167220;40167221;40167222;40167223;40167224;40167225;40167226;40167227;40167228;40167229;40167230;40167231;40167232;40167233;40167234;40167235;40167236;40167237;40167238;40167239;40167240;40167241;40167242;40167243;40167244;40167245;40167246;40167247;40167248;40167249;40167250;40167251;40167252;40167253;40167254;40167255;40167256;40167257;40167258;40167259;40167260;40167261;40167262;40167263;40167264;40167265;40167266;40167267;40167268;40167269;40167270;40167271;40167272;40167273;40167274;40167275;40167276;40167277;40167278;40167279;40167280;40167281;40167282;40167283;40167284;40167285;40167286;40167287;40167288;40167289;40167290;40167291;40167292;40167293;40167294;40167295;40167296;40167297;40167298;40167299;40167300;40167301;40167302;40167303;40167304;40167305;40167306;40167307;40167308;40167309;40167310;40167311;40167312;40167313;40167314;40167315;40167316;40167317;40167318;40167319;40167320;4016
```

Appendix E: Results example treatment patterns study

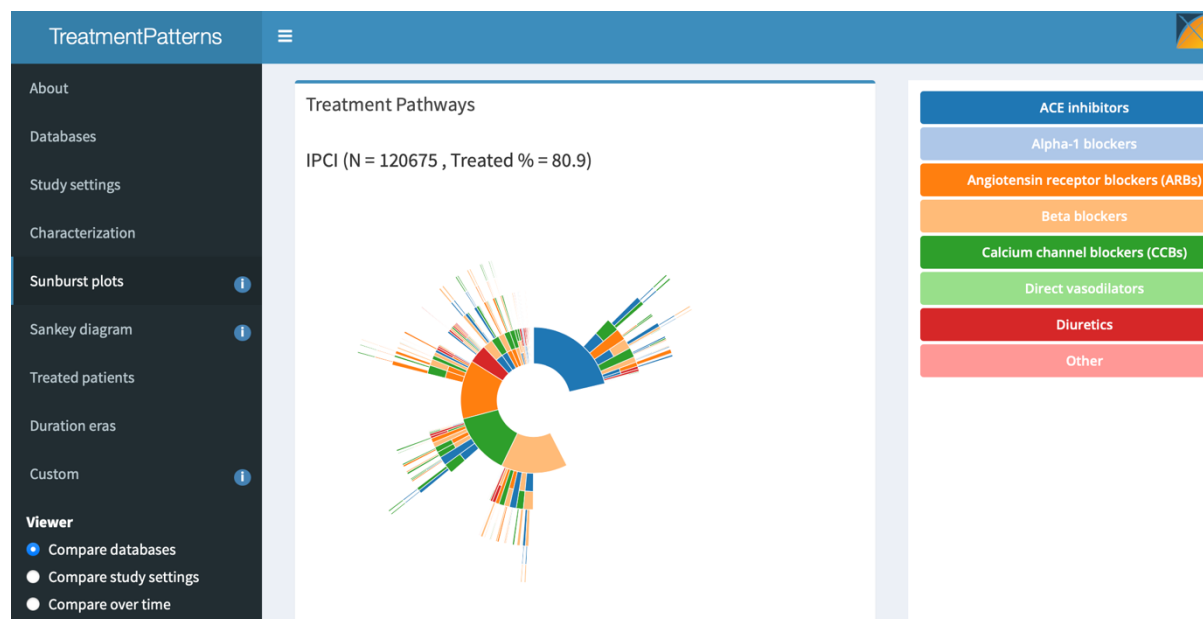


Figure 1: Sunburst plot visualizing the treatment pathways of patients with hypertension in IPCI.

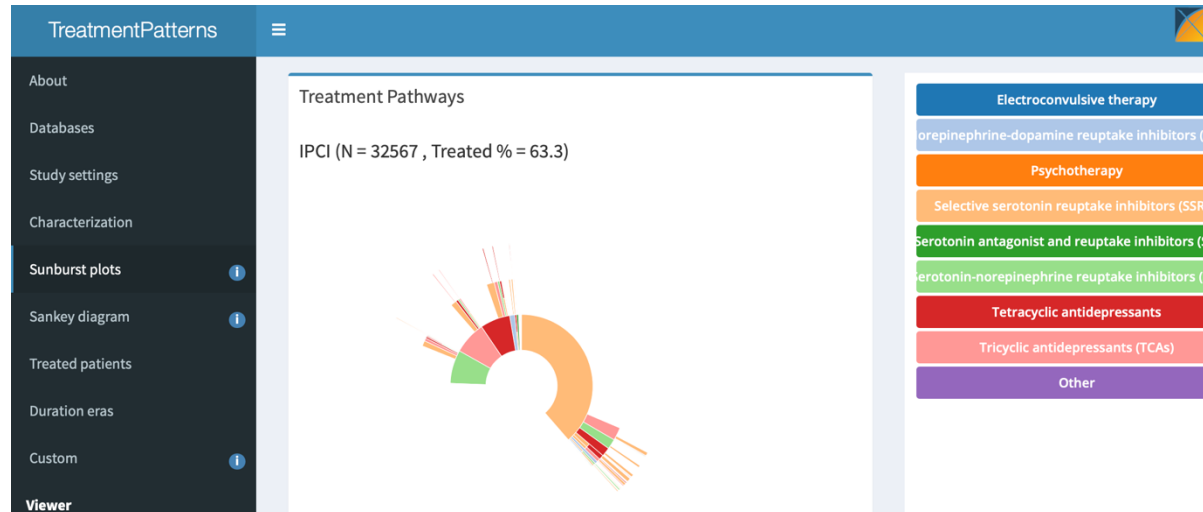


Figure 2: Sunburst plot visualizing the treatment pathways of patients with depression in IPCI.

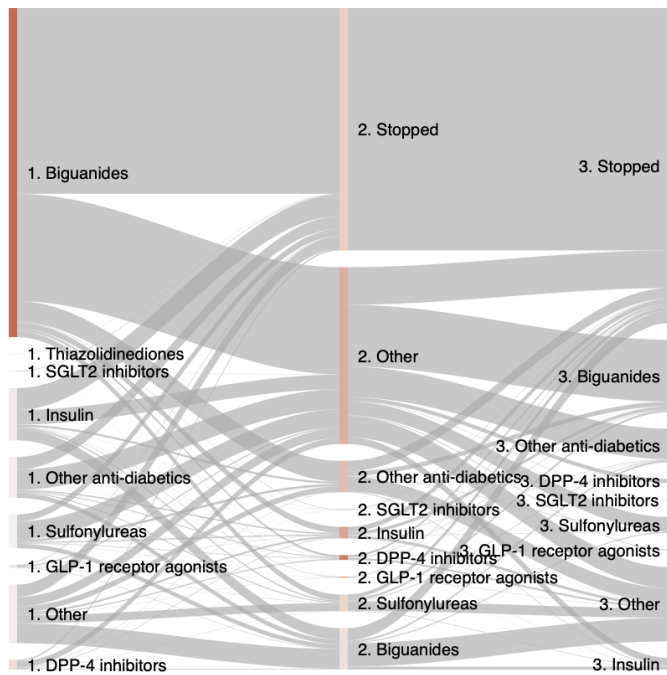


Figure 3: Sankey diagram visualizing the treatment pathways of patients with type II diabetes mellitus in IPCI.

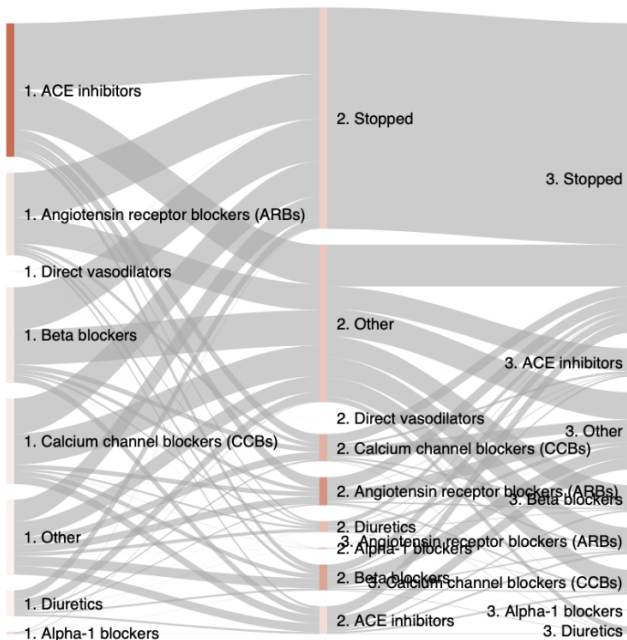


Figure 4: Sankey diagram visualizing the treatment pathways of patients with hypertension in IPCI.

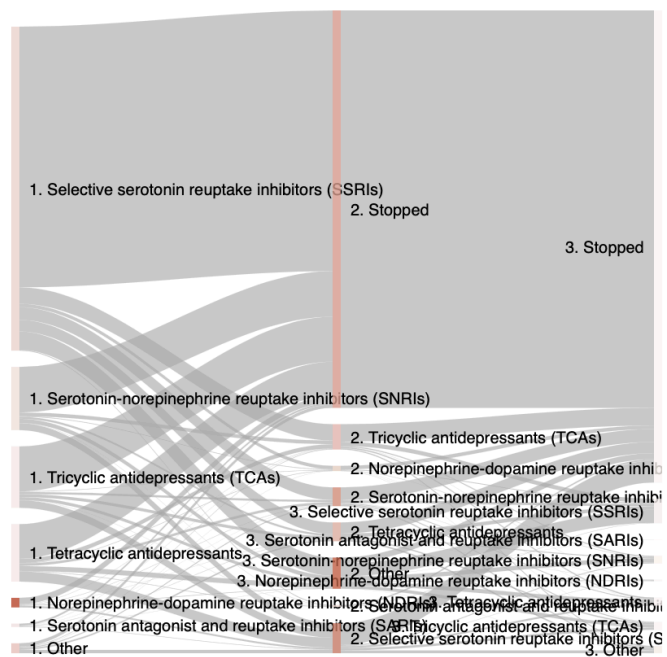


Figure 5: Sankey diagram visualizing the treatment pathways of patients with depression in IPCI.

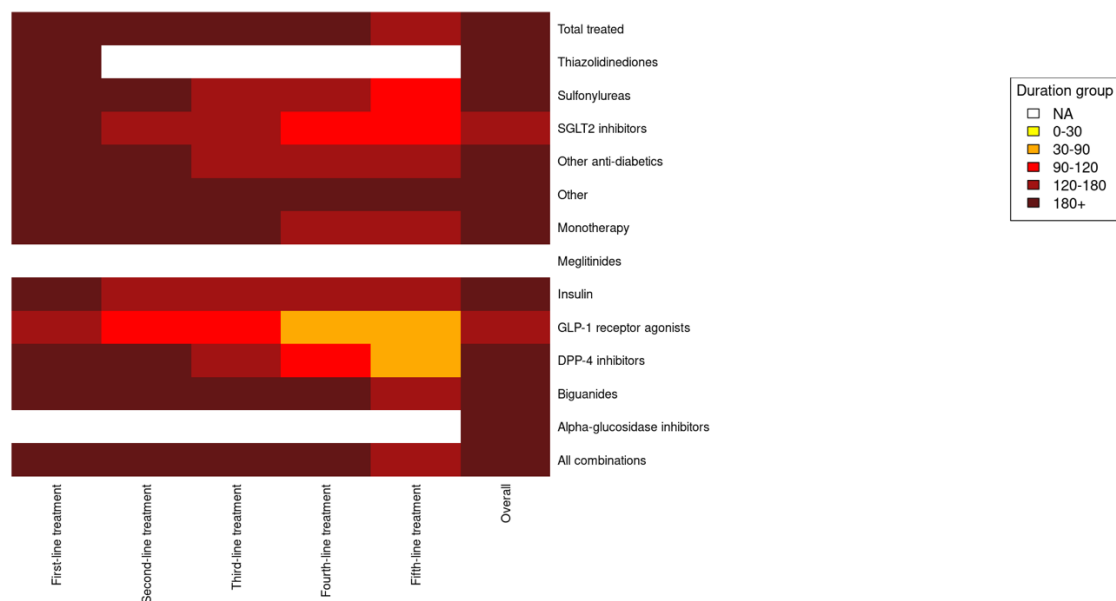


Figure 6: Heatmap visualizing duration of treatments of patients with type II diabetes mellitus in IPCI.

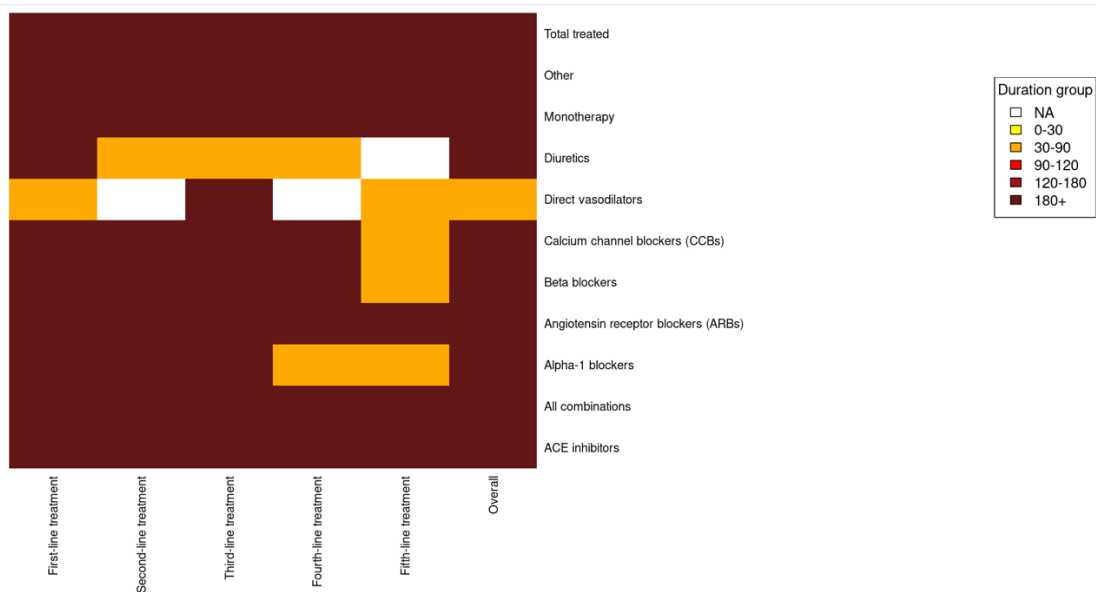


Figure 7: Heatmap visualizing duration of treatments of patients with hypertension in IPCI.

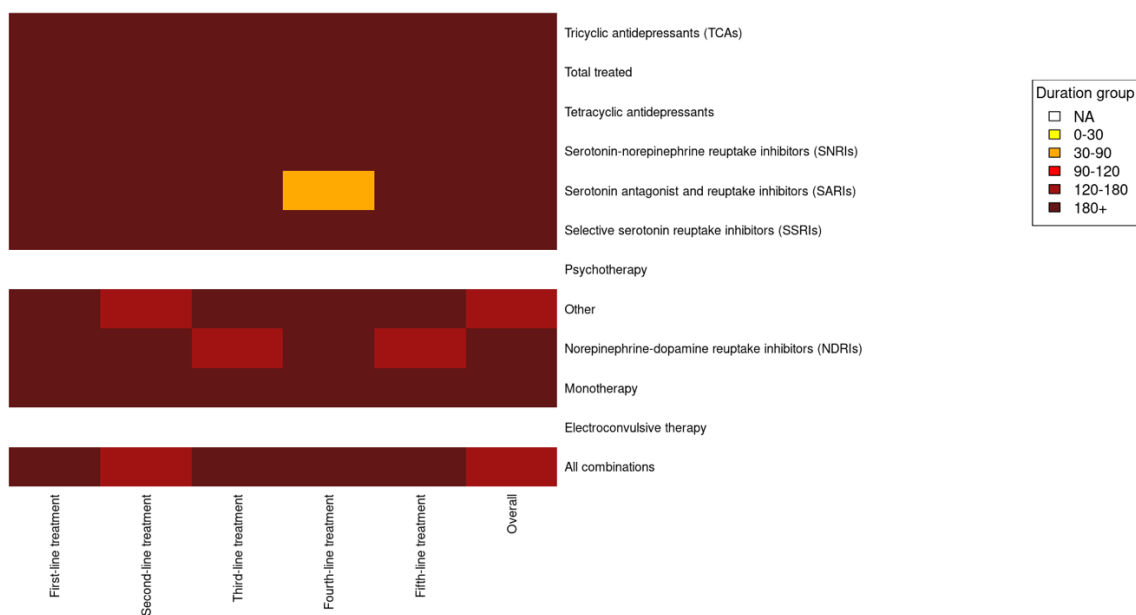


Figure 8: Heatmap visualizing duration of treatments of patients with depression in IPCI.