

RESEARCH

Open Access



# Prescription pattern analysis of Type 2 Diabetes Mellitus: a cross-sectional study in Isfahan, Iran

Elnaz Ziad<sup>1</sup>, Somayeh Sadat<sup>2\*</sup>, Farshad Farzadfar<sup>3</sup> and Mohammad-Reza Malekpour<sup>3</sup>

\*Correspondence:  
Somayeh.sadat@utoronto.ca

<sup>1</sup> School of Industrial and Systems Engineering, Tarbiat Modares University, Tehran, Islamic Republic of Iran

<sup>2</sup> Centre for Analytics and Artificial Intelligence Engineering, University of Toronto, Toronto, Canada

<sup>3</sup> Non-Communicable Diseases Research Center, Endocrinology and Metabolism Population Sciences Institute, Tehran University of Medical Sciences, Tehran, Islamic Republic of Iran

## Abstract

**Background:** Patients with Type 2 Diabetes Mellitus (T2DM) are at a higher risk of polypharmacy and more susceptible to irrational prescriptions; therefore, pharmacological therapy patterns are important to be monitored. The primary objective of this study was to highlight current prescription patterns in T2DM patients and compare them with existing Standards of Medical Care in Diabetes. The second objective was to analyze whether age and gender affect prescription patterns.

**Method:** This cross-sectional study was conducted using the Iran Health Insurance Organization (IHIO) prescription database. It was mined by an Association Rule Mining (ARM) technique, FP-Growth, in order to find co-prescribed drugs with anti-diabetic medications. The algorithm was implemented at different levels of the Anatomical Therapeutic Chemical (ATC) classification system, which assigns different codes to drugs based on their anatomy, pharmacological, therapeutic, and chemical properties to provide an in-depth analysis of co-prescription patterns.

**Results:** Altogether, the prescriptions of 914,652 patients were analyzed, of whom 91,505 were found to have diabetes. According to our results, prescribing Lipid Modifying Agents (C10) (56.3%), Agents Acting on The Renin-Angiotensin System (C09) (48.9%), Antithrombotic Agents (B01) (35.7%), and Beta Blocking Agents (C07) (30.1%) were meaningfully associated with the prescription of Drugs Used in Diabetes. Our study also revealed that female diabetic patients have a higher lift for taking Thyroid Preparations, and the older the patients were, the more they were prone to take neuropathy-related medications. Additionally, the results suggest that there are gender differences in the association between aspirin and diabetes drugs, with the differences becoming less pronounced in old age.

**Conclusions:** Almost all of the association rules found in this research were clinically meaningful, proving the potential of ARM for co-prescription pattern discovery. Moreover, implementing level-based ARM was effective in detecting difficult-to-spot rules. Additionally, the majority of drugs prescribed by physicians were consistent with the Standards of Medical Care in Diabetes.

**Keywords:** Diabetes, Prescription pattern, Association Rule Mining, Multimorbidity



## Background

The WHO Global Report on Diabetes indicates that the number of Diabetes Mellitus (DM) patients has increased globally in recent decades, from 108 million in 1980 [1] to 537 million in 2021 [2]. This almost fivefold increase was mainly due to the rise in Type 2 Diabetes Mellitus (T2DM) and its risk factors, including obesity, overweight, and ageing [1]. If no proper action is taken to prevent this rise, it has been predicted that 783 million individuals will have diabetes by 2045 (12.2% of the total population) [2]. According to the findings of the national STEPwise Approach to NCD Risk Factor Surveillance (STEPS) 2016, the prevalence of diabetes based on HbA1C in Iran was estimated at 11.0%, 12.7%, and 11.9% for males, females, and both genders, respectively [3].

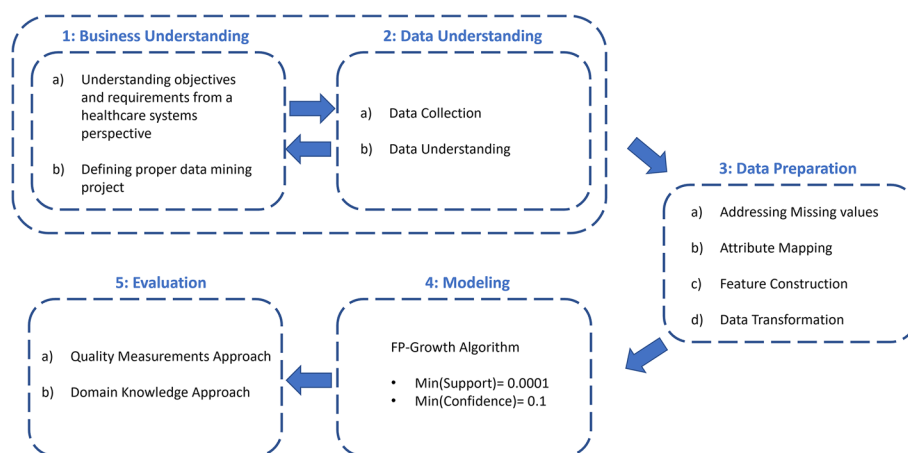
Regarding the growing epidemic of diabetes [4] and its importance as a global public health problem [5], researchers conducted a variety of in-depth studies to analyze different aspects of this long-lasting disease. Various tools and techniques were employed in diabetes research, the most important of which is data mining. Data mining is used to discover unsuspected relationships, identify patterns, and uncover knowledge from raw data [6]. Some diabetes related applications of data mining methods are as follows: diagnosis and prediction of DM [7], effect of genetic background and environment in the development of diabetes [8], adherence to the clinical guidelines for DM [9], diabetic complications [10], prevention of DM [11], medication recommendation [12], biomarkers identification [13], and patterns of pharmacological therapy [14, 15].

Patterns of pharmacological therapy can be mined by several approaches, one of which is discovering co-prescription patterns to analyze multimorbidity. Multimorbidity, defined as the coexistence of two or more health conditions, poses a significant challenge to public health [16]. Multimorbidity often leads to the associated use of multiple medicines, which is known as polypharmacy [17]. Inappropriate polypharmacy increases the potential for medication nonadherence, drug-drug interactions, and adverse drug events; therefore, analysis of polypharmacy in order to provide valuable insight into current co-prescription patterns has become an interesting discipline [18].

The first objective of this study was to identify the drugs prescribed in combination with Drugs Used in Diabetes for T2DM patients in Isfahan, Iran using their insurance claims data and compare the results with the existing Standards of Medical Care in Diabetes. As a second objective, we investigated the effects of age and gender on prescription patterns. Association Rule Mining (ARM) was the most suitable method for the purpose of this study. ARM is widely used and popular among data mining methods, and aims to find interesting rules from transactional datasets. The main novelty of this study is providing a detailed drill-down analysis of co-prescription patterns for T2DM patients based on different levels defined by the Anatomical Therapeutic Chemical (ATC) coding system.

## Materials and methods

The data analysis pipeline of this research conforms to the Cross-Industry Standard Process for Data Mining (CRISP-DM) methodology [19] and is depicted in Fig. 1. It contains five phases, including Business Understanding, Data Understanding, Data Preparation, Modeling, and Evaluation.



**Fig. 1** Data analysis pipeline

### Phase 1 and 2: business and data understanding

There is a close link between phase 1 and phase 2 since formulating the data mining problem requires at least some understanding of the available data.

The datasets used in this study were acquired from the Iran Health Insurance Organization (IHIO). The following tables were retrieved from the IHIO data warehouse:

- 1) Electronic insurance claims fact table, which contained fields for drug prescription information, patient demographics, and the prescriber's medical council's code. This table contained the prescriptions of 914,652 unique patients.
- 2) Dimension tables, first of which contained national drug codes as keys and different levels of ATC drug codes as values, and the second of which had the medical council's codes as key and the specialty of the physicians as values.

We also needed to understand the study requirements and objectives from a healthcare systems perspective, as well as define a proper data mining project to achieve research objectives.

Insurance claims data tend to be incomplete, and one of the major requirements for any prescription-based study is the diagnosis label. The data used in this study lack the International Classification of Disease (ICD) code for each prescription, which is why the first challenge is to answer this question: "How to identify patients with T2DM using insurance claims data accurately?" Therefore, a rule framework was implemented based on some clinical assumptions to differentiate between T2DM patients taking drugs from the A10 subgroup and patients taking them for other reasons such as Polycystic Ovary Syndrome (PCOS), Type 1 Diabetes Mellitus (T1DM), or pregnancy. A panel of five medical experts defined the following criteria to recognize T2DM in subjects:

1. Presence of antidiabetic drugs in the prescription:
  - a. Oral Anti Diabetic agents (OADs)
  - b. Insulin

Antidiabetic drugs can be determined by means of the ATC classification system. In this system, the active substances are divided into different groups according to the organ or system on which they act and their therapeutic, pharmacological, and chemical properties [20]. A10 is a predefined level of ATC, indicating Drugs Used in Diabetes. If a prescription contains at least one of the A10 drugs, there is a possibility that the patient suffers from T2DM. However, it is not the sole criterion for labeling T2DM patients.

2. The specialty of prescribing physician:

- a. General practitioner
- b. Internist
- c. Endocrinologist

If the prescriber's specialty is one of the mentioned specialties, it raises the possibility that the prescription belongs to a T2DM patient.

3. Checking for other conditions

- a. T1DM in males (if the prescription belongs to a male patient under the age of 30 and contains any kind of Insulin)
- b. T1DM in females (if the prescription belongs to a female patient under the age of 30 and contains any kind of Insulin except Insulin Human)
- c. PCOS (if the prescription belongs to a female patient under the age of 30 and contains metformin prescribed by a gynecology physician)
- d. Pregnancy (if the prescription is for a female patient under the age of 30, contains Insulin Human, and the patient is not on the list of patients who have taken at least one OAD in a non-PCOS prescription)

Subjects who met the criteria in previous steps may also have T1DM, PCOS, or be pregnant. Therefore, it is crucial to draw a distinction between T2DM and other conditions.

To sum up, when a patient takes one of the A10 subgroup drugs (excluding Insulin siring), the prescriber has one of the mentioned specialties, and the patient does not suffer from other conditions, the patient can be labeled as a T2DM patient. In this study, 91,505 patients were labeled as having T2DM, and A10 drugs were removed from the prescriptions of patients who were not labeled as having T2DM.

To convert the objective of this study to a data mining problem definition, the ARM algorithm was selected in this phase. Several ARM algorithms were developed, including Apriori, Eclat, and FP-Growth. In this study, FP-growth was selected because it is faster, more efficient, more scalable than other algorithms, and it is used in recently published articles [21, 22].

**Phase 3: data preprocessing**

Apache Spark [23], version 3.2.0, a unified big data processing engine, was utilized for data preprocessing. A brief description of this procedure is provided in the subsections below.

- **Addressing missing values**

The patients with missing gender and age are just a random subset of all patients, so there are no meaningful differences between these patients and the others. It implies a Missing Completely At Random (MCAR) situation, in which it is safe to remove rows with missing values because the results will be unbiased. Therefore, missing values were handled by deleting the entire prescriptions of patients having NULL gender and age.

- **Attribute mapping**

Attribute mapping is the act of transforming and/or connecting one or more attributes to a new attribute or set of attributes. In this step, attribute mapping was used as a solution to the following problem:

Drug codes were based on the national Food and Drug Administration coding system in the insurance claims dataset. It was necessary to transform national drug codes into international ones that are based on the ATC Classification System. First, a dataset that contained the corresponding ATC codes for every national drug code was retrieved. Then, attribute mapping was carried out to retrieve different levels of ATC codes in separate columns (Table 1). The same action was taken to map the medical counsel’s code of prescribers to their specialty and the age of the patient to the matching age category (Tables 2 and 3). In this study, three age categories were defined as follows:

- Age < 30 -> Young
- 30 <= Age < 45 -> Middle-Aged
- Age >= 45 -> Old

**Table 1** Attribute mapping for national drug ID

...	National_Drug_ID	Generic_Drug_Name	ATC_L5	ATC_L4	ATC_L3	ATC_L2	ATC_L1
...	02023	Asa 100 mg - tablet	B01AC06	B01AC	B01A	B01	B

**Table 2** Attribute mapping for prescriber medical council code

...	Prescriber_Medical_Council_Code	Prescriber Specialty
...	1787	Endocrinologist

**Table 3** Attributing mapping for age

...	Age	Age_Category
...	38	Middle-Aged

- **Feature construction**

Patients’ national codes, which were unquestionably unique for each individual, were not available due to data anonymization. Therefore, creating a unique ID for each patient was necessary. In the available dataset, Insurance\_ID was not unique. Family members had the same Insurance\_ID but different Insurance\_Serials. To create a unique ID for each patient, the Insurance\_ID and Insurance\_Serial features were combined and named Patient\_ID.

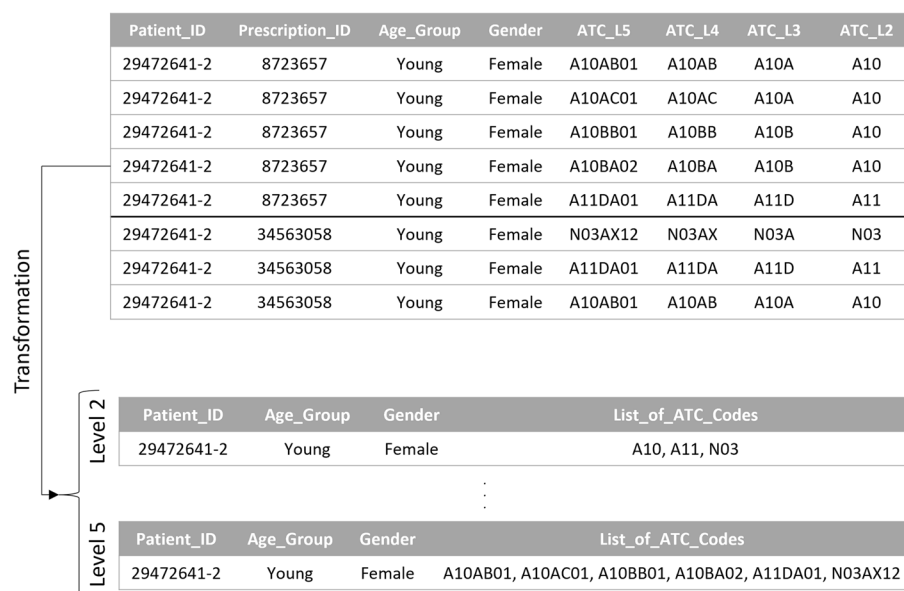
- **Data transformation**

This step involved data transformation so it could be used as input for the FP-Growth algorithm. First, the data were aggregated at the patient level for each level of ATC. Second, duplicate ATC codes were removed to create a unique list. This procedure was done for all combinations of age categories and genders. In Fig. 2, an example of this transformation is provided.

**Phase 4: modeling**

FP-growth is currently one of the fastest algorithms for finding frequent patterns in transactional databases due to its independence from candidate generation and ability to store a compact version of the database in memory [21]. In the first step, this algorithm finds frequent items by calculating their frequencies. In the second step, it uses a suffix tree (FP-tree) structure to encode transactions without explicitly generating candidate sets. Finally, the frequent itemsets can be extracted from the FP-tree. In this study, FP-growth was implemented in Apache Spark [23] using the FPGrowthModel.

Another indicator, named prevalence, was also computed for the extracted rules of the second level of ATC. Let us denote Drugs Used in Diabetes as A and another drug as B. Regarding the formula for support,  $Supp(A \rightarrow B) = \sigma(A \rightarrow B)/N$ , and knowing the value of N, 914,652, we can calculate  $\sigma(A \rightarrow B)$ , which shows the number of patients who took



**Fig. 2** Data transformation

drugs A and B. In this section, we calculated the proportion of  $\sigma(A \rightarrow B)$  to only diabetic patients, 91,505, to compute the prevalence of prescribing drug B among T2DM patients.

#### Phase 5: evaluation

In this phase, the details of the utilized validation approaches will be explained.

a) **Quality measurements approach:** Not all the rules generated by ARM algorithms are meaningful. Quality measurements can be helpful in excluding some coincidental rules. Each Association Rule (AR) has two primary quality measurements, namely support and confidence, defined as:

- **Support criterion:** the support of an AR denoted by  $\text{Supp}(A \rightarrow B)$  is defined as  $\text{Supp}(A \rightarrow B) = \sigma(A \rightarrow B)/N$ , where  $\sigma(A \rightarrow B)$  is the number of patients who have taken drugs A and B, and N is the total number of patients. In sum, support is simply the frequency of occurrence of each rule.
- **Confidence criterion:** the confidence of an AR denoted by  $\text{Conf}(A \rightarrow B)$  is defined as  $\text{Conf}(A \rightarrow B) = (\text{Supp}(A \rightarrow B)) / (\text{Supp}(A))$ , where  $\text{Supp}(A)$  is the number of patients who have taken drug A. In summary, confidence is the strength of implication.

A rule is considered valid if it satisfies the below conditions:

$$\begin{aligned} \text{Support}(A \rightarrow B) &\geq \text{Minisupp} \\ \text{Confidence}(A \rightarrow B) &\geq \text{Miniconf} \end{aligned}$$

In this study, we performed a grid search and opted for  $\text{support} = 0.0001$  and  $\text{confidence} = 0.1$  to uncover valuable associations across different patient segments while avoiding redundancy.

b) **Domain knowledge approach:** after passing the previous step, two medical doctors investigated the remaining rules in order to check clinical rationality and possible indications.

## Results

Overall, we analyzed the prescriptions of 914,652 unique patients, of whom 91,505 were diabetic. It suggests that the prevalence of T2DM in our sample was approximately 10%. In total population, there were 510,873 female and 403,779 male patients; these figures were 58,023 and 33,482 for diabetic patients. Regarding age groups, there were 355,925 patients in the “young” category, followed by 203,166 and 355,561 in the “middle-aged” and “old” categories. The mentioned figures were 6,750, 14,092, and 70,663 in the diabetic population, respectively. The average age of the total population was 38.67, with a standard deviation of 21.78, while diabetic patients were on average 56.08 years old, with a standard deviation of 15.83. In the following subsections, we decided to present the top 30 rules of each ATC level with the condition of lift value above 2.0; therefore, some levels had less than 30 ARs regarding this criterion.

**Second level of ATC**

ARs were sorted by prevalence indicator since we wanted to present the most prescribed drug classes (Table 4). According to this indicator, Lipid Modifying Agents (C10) with 56.3%, Agents Acting on The Renin-Angiotensin System (C09) with 48.9%, Antithrombotic Agents (B01) with 35.7%, and Beta Blocking Agents (C07) with 30.1% were the most prevalent drug classes prescribed with Drugs Used in Diabetes (A10).

**Third level of ATC**

Patients who took both Insulins and Analogues (A10A) and Blood Glucose Lowering Drugs, Excl. Insulins (A10B) were more likely to take ACE Inhibitors, Plain (C09A) by 5.43 times, followed by Vitamin B1, Plain and In Combination with Vitamin B6 And B12 (A11D; 5.39), Angiotensin II Receptor Blockers (ARBs), Plain (C09C; 5.08), Lipid Modifying Agents, Plain (C10A; 4.87), and Antithrombotic Agents (B01A; 4.17) (Table 5).

**Fourth level of ATC**

The top rule regarding lift value was “Insulins and Analogues for Injection, Intermediate-Acting, Biguanides, Sulfonylureas = > Ace Inhibitors, Plain”, with a lift of 10.13 (Table 6). Another meaningful rule had the same LHS as the previously mentioned top rule, and its RHS is Vitamin B1, Plain, with a lift of 9.04. The ARs number 7, 9, 14, 19, and 20 suggest that among C03C, C03A, C01D, C10A, and B01A subsets, Sulfonamides, Plain (C03CA), Thiazides, Plain (C03AA), Organic Nitrates (C01DA), HMG-CoA Reductase Inhibitors (C10AA), and Platelet Aggregation Inhibitors Excl. Heparin (B01AA) were the main choices of physicians, with lift values of 7.66, 6.59, 6.28, 6.28, and 6.06, respectively.

**Table 4** ARs of the 2<sup>nd</sup> ATC level

Num	LHS	RHS	LHS name	RHS name	Conf	Lift	Supp	Prevalence
1	[A10]	[C10]	Drugs Used In Diabetes	Lipid Modifying Agents	0.4804	3.7129	0.0563	56.3%
2	[A10]	[C09]	Drugs Used In Diabetes	Agents Acting On The Renin-Angiotensin System	0.4179	3.4548	0.0489	48.9%
3	[A10]	[B01]	Drugs Used In Diabetes	Antithrombotic Agents	0.3045	2.9911	0.0357	35.7%
4	[A10]	[C07]	Drugs Used In Diabetes	Beta Blocking Agents	0.2570	2.1713	0.0301	30.1%
5	[A10]	[C03]	Drugs Used In Diabetes	Diuretics	0.1479	2.8212	0.0173	17.3%
6	[A10]	[C08]	Drugs Used In Diabetes	Calcium Channel Blockers	0.1478	2.9207	0.0173	17.3%
7	[A10]	[C01]	Drugs Used In Diabetes	Cardiac Therapy	0.1342	2.5290	0.0157	15.7%

**Table 5** ARs of the 3<sup>rd</sup> ATC level

Num	LHS	RHS	LHS name	RHS name	Conf	Lift	Supp
1	[A10A, A10B]	[C09A]	Insulins And Analogues, Blood Glucose Lowering Drugs, Excl. Insulins	Ace Inhibitors, Plain	0.1124	5.4306	0.0016
2	[A10A, A10B]	[A11D]	Insulins And Analogues, Blood Glucose Lowering Drugs, Excl. Insulins	Vitamin B1, Plain And In Combination With Vitamin B6 And B12	0.1100	5.3923	0.0016
3	[A10A, A10B]	[C09C]	Insulins And Analogues, Blood Glucose Lowering Drugs, Excl. Insulins	Angiotensin II Receptor Blockers (Arbs), Plain	0.5201	5.0870	0.0076
4	[A10A, A10B]	[C10A]	Insulins And Analogues, Blood Glucose Lowering Drugs, Excl. Insulins	Lipid Modifying Agents, Plain	0.6310	4.8763	0.0092
5	[A10A, A10B]	[B01A]	Insulins And Analogues, Blood Glucose Lowering Drugs, Excl. Insulins	Antithrombotic Agents	0.4258	4.1794	0.0062
6	[A10A, A10B]	[C03A]	Insulins And Analogues, Blood Glucose Lowering Drugs, Excl. Insulins	Low-Ceiling Diuretics, Thiazides	0.1345	4.0818	0.0020
7	[A10B]	[C10A]	Blood Glucose Lowering Drugs, Excl. Insulins	Lipid Modifying Agents, Plain	0.5139	3.9713	0.0503
8	[A10A]	[C09C]	Insulins And Analogues	Angiotensin II Receptor Blockers (Arbs), Plain	0.3853	3.7685	0.0130
9	[A10A, A10B]	[C01D]	Insulins And Analogues, Blood Glucose Lowering Drugs, Excl. Insulins	Vasodilators Used In Cardiac Diseases	0.1805	3.7191	0.0026
10	[A10A, A10B]	[C08C]	Insulins And Analogues, Blood Glucose Lowering Drugs, Excl. Insulins	Selective Calcium Channel Blockers With Mainly Vascular Effects	0.1447	3.6749	0.0021
11	[A10B]	[C09C]	Blood Glucose Lowering Drugs, Excl. Insulins	Angiotensin II Receptor Blockers (Arbs), Plain	0.3678	3.5971	0.0360
12	[A10A]	[C10A]	Insulins And Analogues	Lipid Modifying Agents, Plain	0.4486	3.4669	0.0152
13	[A10A]	[C01D]	Insulins And Analogues	Vasodilators Used In Cardiac Diseases	0.1591	3.2795	0.0054
14	[A10A]	[B01A]	Insulins And Analogues	Antithrombotic Agents	0.3240	3.1825	0.0109
15	[A10B]	[B01A]	Blood Glucose Lowering Drugs, Excl. Insulins	Antithrombotic Agents	0.3158	3.1022	0.0309
16	[A10B]	[C03A]	Blood Glucose Lowering Drugs, Excl. Insulins	Low-Ceiling Diuretics, Thiazides	0.1006	3.0539	0.0099
17	[A10B]	[C08C]	Blood Glucose Lowering Drugs, Excl. Insulins	Selective Calcium Channel Blockers With Mainly Vascular Effects	0.1199	3.0451	0.0117
18	[A10A]	[C08C]	Insulins And Analogues	Selective Calcium Channel Blockers With Mainly Vascular Effects	0.1194	3.0322	0.0040
19	[A10A, A10B]	[C07A]	Insulins And Analogues, Blood Glucose Lowering Drugs, Excl. Insulins	Beta Blocking Agents	0.3130	2.6442	0.0046
20	[A10B]	[C01D]	Blood Glucose Lowering Drugs, Excl. Insulins	Vasodilators Used In Cardiac Diseases	0.1236	2.5476	0.0121
21	[A10B]	[C07A]	Blood Glucose Lowering Drugs, Excl. Insulins	Beta Blocking Agents	0.2683	2.2667	0.0263
22	[A10A]	[C07A]	Insulins And Analogues	Beta Blocking Agents	0.2485	2.0991	0.0084



**Table 6** ARs of the 4<sup>th</sup> ATC level

Num	LHS	RHS	LHS name	RHS name	Conf	Lift	Supp
1	[A10AC, A10BB, A10BA]	[C09AA]	Insulins And Analogues For Injection, Intermediate-Acting, Biguanides, Sulfonylureas	Ace Inhibitors, Plain	0.2099	10.1359	0.0001
2	[A10AC, A10BB]	[C09AA]	Insulins And Analogues For Injection, Intermediate-Acting, Sulfonylureas	Ace Inhibitors, Plain	0.2011	9.7100	0.0002
3	[A10AD, A10BB, A10BA]	[A11DA]	Insulins And Analogues For Injection, Intermediate- Or Long-Acting Combined With Fast-Acting, Biguanides, Sulfonylureas	Vitamin B1, Plain	0.1847	9.0492	0.0001
4	[A10AD, A10BB]	[A11DA]	Insulins And Analogues For Injection, Intermediate- Or Long-Acting Combined With Fast-Acting, Sulfonylureas	Vitamin B1, Plain	0.1700	8.3304	0.0001
5	[A10AB, A10BB]	[C09AA]	Insulins And Analogues For Injection, Fast-Acting, Sulfonylureas	Ace Inhibitors, Plain	0.1646	7.9509	0.0001
6	[A10AD, A10AE]	[A11DA]	Insulins And Analogues For Injection, Intermediate- Or Long-Acting Combined With Fast-Acting, Insulins And Analogues For Injection, Long-Acting	Vitamin B1, Plain	0.1601	7.8468	0.0001
7	[A10AC, A10AB]	[C03CA]	Insulins And Analogues For Injection, Fast-Acting, Insulins And Analogues For Injection, Intermediate-Acting	Sulfonamides, Plain	0.1243	7.6603	0.0012
8	[A10AC]	[C03CA]	Insulins And Analogues For Injection, Intermediate-Acting	Sulfonamides, Plain	0.1192	7.3451	0.0016
9	[A10AD, A10AC, A10AB, A10BA]	[C09CA]	Insulins And Analogues For Injection, Fast-Acting, Insulins And Analogues For Injection, Intermediate-Acting, Insulins And Analogues For Injection, Intermediate- Or Long-Acting Combined With Fast-Acting, Biguanides	Angiotensin II Receptor Blockers (Arbs), Plain	0.7143	6.9865	0.0001
10	[A10AC, A10AB, A10BA]	[C09AA]	Insulins And Analogues For Injection, Fast-Acting, Insulins And Analogues For Injection, Intermediate-Acting, Biguanides	Ace Inhibitors, Plain	0.1431	6.9117	0.0005
11	[A10AE, A10AB, A10BA]	[A11DA]	Insulins And Analogues For Injection, Fast-Acting, Insulins And Analogues For Injection, Long-Acting, Biguanides	Vitamin B1, Plain	0.1401	6.8675	0.0002
12	[A10AC, A10BA]	[C09AA]	Insulins And Analogues For Injection, Intermediate-Acting, Biguanides	Ace Inhibitors, Plain	0.1414	6.8302	0.0006
13	[A10AB]	[C03CA]	Insulins And Analogues For Injection, Fast-Acting	Sulfonamides, Plain	0.1091	6.7209	0.0016
14	[A10AC, A10AB, A10BA]	[C03AA]	Insulins And Analogues For Injection, Fast-Acting, Insulins And Analogues For Injection, Intermediate-Acting, Biguanides	Thiazides, Plain	0.1467	6.5903	0.0005
15	[A10AD, A10AC, A10BA]	[C09CA]	Insulins And Analogues For Injection, Intermediate-Acting, Insulins And Analogues For Injection, Intermediate- Or Long-Acting Combined With Fast-Acting, Biguanides	Angiotensin II Receptor Blockers (Arbs), Plain	0.6696	6.5498	0.0002
16	[A10AE, A10AB, A10BB, A10BA]	[C10AA]	Insulins And Analogues For Injection, Fast-Acting, Insulins And Analogues For Injection, Long-Acting, Biguanides, Sulfonylureas	Hmg Coa Reductase Inhibitors	0.7910	6.4560	0.0001
17	[A10AC, A10BA]	[C03AA]	Insulins And Analogues For Injection, Intermediate-Acting, Biguanides	Thiazides, Plain	0.1419	6.3722	0.0006
18	[A10AC, A10AB, A10BA]	[C03CA]	Insulins And Analogues For Injection, Fast-Acting, Insulins And Analogues For Injection, Intermediate-Acting, Biguanides	Sulfonamides, Plain	0.1021	6.2873	0.0003
19	[A10AC, A10AB, A10BB]	[C01DA]	Insulins And Analogues For Injection, Fast-Acting, Insulins And Analogues For Injection, Intermediate-Acting, Sulfonylureas	Organic Nitrates	0.3049	6.2837	0.0001
20	[A10BG, A10AD, A10BA]	[C10AA]	Insulins And Analogues For Injection, Intermediate- Or Long-Acting Combined With Fast-Acting, Biguanides, Thiazolidinediones	Hmg Coa Reductase Inhibitors	0.7699	6.2836	0.0002
21	[A10AD, A10AC, A10AB, A10BA]	[C10AA]	Insulins And Analogues For Injection, Fast-Acting, Insulins And Analogues For Injection, Intermediate-Acting, Insulins And Analogues For Injection, Intermediate- Or Long-Acting Combined With Fast-Acting, Biguanides	Hmg Coa Reductase Inhibitors	0.7662	6.2536	0.0001
22	[A10AD, A10BA]	[A11DA]	Insulins And Analogues For Injection, Intermediate- Or Long-Acting Combined With Fast-Acting, Biguanides	Vitamin B1, Plain	0.1271	6.2264	0.0006
23	[A10AC, A10BB]	[C03AA]	Insulins And Analogues For Injection, Intermediate-Acting, Sulfonylureas	Thiazides, Plain	0.1376	6.1783	0.0001
24	[A10BF, A10AD, A10BA]	[C10AA]	Insulins And Analogues For Injection, Intermediate- Or Long-Acting Combined With Fast-Acting, Biguanides, Alpha Glucosidase Inhibitors	Hmg Coa Reductase Inhibitors	0.7536	6.1506	0.0001
25	[A10AD, A10AC, A10AB]	[C09CA]	Insulins And Analogues For Injection, Fast-Acting, Insulins And Analogues For Injection, Intermediate-Acting, Insulins And Analogues For Injection, Intermediate- Or Long-Acting Combined With Fast-Acting	Angiotensin II Receptor Blockers (Arbs), Plain	0.6250	6.1132	0.0002
26	[A10BG, A10AC, A10BA]	[C09CA]	Insulins And Analogues For Injection, Intermediate-Acting, Biguanides, Thiazolidinediones	Angiotensin II Receptor Blockers (Arbs), Plain	0.6250	6.1132	0.0001
27	[A10AE, A10AB, A10BB]	[C10AA]	Insulins And Analogues For Injection, Fast-Acting, Insulins And Analogues For Injection, Long-Acting, Sulfonylureas	Hmg Coa Reductase Inhibitors	0.7474	6.0996	0.0001
28	[A10BG, A10BB, A10BA]	[C09AA]	Biguanides, Sulfonylureas, Thiazolidinediones	Ace Inhibitors, Plain	0.1258	6.0744	0.0005
29	[A10AD, A10AC, A10BA]	[B01AC]	Insulins And Analogues For Injection, Intermediate-Acting, Insulins And Analogues For Injection, Intermediate- Or Long-Acting Combined With Fast-Acting, Biguanides	Platelet Aggregation Inhibitors Excl. Heparin	0.5714	6.0636	0.0001
30	[A10BG, A10AB, A10BA]	[C10AA]	Insulins And Analogues For Injection, Fast-Acting, Biguanides, Thiazolidinediones	Hmg Coa Reductase Inhibitors	0.7407	6.0455	0.0001

**Fifth level of ATC**

According to the results of this level, “Insulin Aspart, Glibenclamide = > Thiamine (Vit B1)” and “Insulin (Human) = > Furosemide” had the highest lift value, with 9.35 and 8.08, respectively (Table 7). Regarding the RHS, ARs associated with Losartan, Atorvastatin, Hydrochlorothiazide, and Gabapentin had the highest lift values, including “Glibenclamide, Gliclazide, Pioglitazone = > Losartan” with a lift of 7.08, “Insulin Aspart, Metformin, Pioglitazone = > Atorvastatin” with a lift of 6.66, “Insulin (Human), Insulin (Human), Metformin = > Hydrochlorothiazide” with a lift of 6.61, and “Insulin Aspart, Insulin Glargine = > Gabapentin” with a lift of 6.44 (Table 7).

**Age and gender-wise results**

The ARs extracted for each combination of gender and age groups are presented in an additional file (see Additional file 1).

**Discussion**

**Second level of ATC**

Although it is not accurate and straightforward to imply the exact prevalence of diabetes complications from the percentage of prescribed drugs, Table 4 may at least approximately suggest the prevalence order of the main diabetes complications. According to medical experts’ opinions, C10, C09, and B01, which are the top three associated medications, are mainly prescribed for treating Dyslipidemia, Hypertension, and Cardiovascular disease, respectively.

**Table 7** ARs of the 5<sup>th</sup> ATC level

Num	LHS	RHS	LHS name	RHS name	Conf	Lift	Supp
1	[A10AD05, A10BB01]	[A11DA01]	Insulin Aspart, Glibenclamide	Thiamine (Vit B1)	0.1908	9.3516	0.0001
2	[A10AD01]	[C03CA01]	Insulin (Human)	Furosemide	0.1313	8.0899	0.0002
3	[A10AD05, A10AE04]	[A11DA01]	Insulin Aspart, Insulin Glargine	Thiamine (Vit B1)	0.1639	8.3005	0.0001
4	[A10AB01, A10AC01]	[C03CA01]	Insulin (Human), Insulin (Human)	Furosemide	0.1248	7.6911	0.0012
5	[A10AB01]	[C03CA01]	Insulin (Human)	Furosemide	0.1229	7.5694	0.0012
6	[A10AC01]	[C03CA01]	Insulin (Human)	Furosemide	0.1193	7.3474	0.0016
7	[A10BF01, A10BB09, A10BA02]	[A11DA01]	Metformin, Gliclazide, Acarbose	Thiamine (Vit B1)	0.1484	7.2704	0.0001
8	[A10BB09, A10BG03, A10BB01]	[C09CA01]	Glibenclamide, Gliclazide, Pioglitazone	Losartan	0.6535	7.0877	0.0002
9	[A10BF01, A10BB09, A10BG03, A10BA02]	[C09CA01]	Metformin, Gliclazide, Acarbose, Pioglitazone	Losartan	0.6484	7.0314	0.0001
10	[A10BB09, A10BG03, A10BB01, A10BA02]	[C09CA01]	Metformin, Glibenclamide, Gliclazide, Pioglitazone	Losartan	0.6466	7.0119	0.0002
11	[A10BF01, A10BB09]	[A11DA01]	Gliclazide, Acarbose	Thiamine (Vit B1)	0.1418	6.9462	0.0001
12	[A10AB05, A10AE04, A10BA02]	[A11DA01]	Insulin Aspart, Insulin Glargine, Metformin	Thiamine (Vit B1)	0.1376	6.7434	0.0002
13	[A10AE04, A10BB01, A10BA02]	[A11DA01]	Insulin Glargine, Metformin, Glibenclamide	Thiamine (Vit B1)	0.1372	6.7233	0.0001
14	[A10BG03, A10AD05, A10BA02]	[C10AA05]	Insulin Aspart, Metformin, Pioglitazone	Atorvastatin	0.7736	6.6619	0.0002
15	[A10AB01, A10AC01, A10BA02]	[C03AA03]	Insulin (Human), Insulin (Human), Metformin	Hydrochlorothiazide	0.1474	6.6183	0.0005
16	[A10AB01, A10BA02]	[C03AA03]	Insulin (Human), Metformin	Hydrochlorothiazide	0.1464	6.5757	0.0005
17	[A10AD05, A10AC01, A10BA02]	[C09CA01]	Insulin (Human), Insulin Aspart, Metformin	Losartan	0.6061	6.5728	0.0001
18	[A10AD05, A10AE04]	[N03AX12]	Insulin Aspart, Insulin Glargine	Gabapentin	0.1906	6.4422	0.0001
19	[A10BB09, A10BG03, A10BB01]	[C10AA05]	Glibenclamide, Gliclazide, Pioglitazone	Atorvastatin	0.7480	6.4419	0.0002
20	[A10AE04, A10BB01]	[A11DA01]	Insulin Glargine, Glibenclamide	Thiamine (Vit B1)	0.1310	6.4182	0.0001
21	[A10AB05, A10BA02]	[A11DA01]	Insulin Aspart, Metformin	Thiamine (Vit B1)	0.1309	6.4161	0.0002
22	[A10AD05, A10AB01, A10AC01, A10BA02]	[C10AA05]	Insulin (Human), Insulin (Human), Insulin Aspart, Metformin	Atorvastatin	0.7432	6.4006	0.0001
23	[A10BB09, A10BG03, A10BB01, A10BA02]	[C10AA05]	Metformin, Glibenclamide, Gliclazide, Pioglitazone	Atorvastatin	0.7414	6.3846	0.0002
24	[A10AC01, A10BA02]	[C03AA03]	Insulin (Human), Metformin	Hydrochlorothiazide	0.1419	6.3722	0.0006
25	[A10BB09, A10BB01, A10BA02]	[C03AA03]	Metformin, Glibenclamide, Gliclazide	Hydrochlorothiazide	0.1418	6.3672	0.0001
26	[A10BB09, A10BB01]	[C09CA01]	Glibenclamide, Gliclazide	Losartan	0.5845	6.3387	0.0005
27	[A10AD05, A10BB01]	[N03AX12]	Insulin Aspart, Glibenclamide	Gabapentin	0.1870	6.3201	0.0001
28	[A10AD05, A10AB01, A10BA02]	[C10AA05]	Insulin (Human), Insulin Aspart, Metformin	Atorvastatin	0.7333	6.3153	0.0001
29	[A10AB01, A10AC01, A10BA02]	[C03CA01]	Insulin (Human), Insulin (Human), Metformin	Furosemide	0.1025	6.3140	0.0003
30	[A10BF01, A10BB09, A10BA02]	[C09CA01]	Metformin, Gliclazide, Acarbose	Losartan	0.5816	6.3075	0.0004

**Third level of ATC**

The results of ARM analysis of the third level of ATC (Table 5) showed that ARs with Right-Hand-Side (RHS) of ACE Inhibitors (C09A) or Angiotensin II Receptor Blockers (ARBs) (C09C) had significant values for the lift indicator. There is a clear recommendation in the Standards of Medical Care in Diabetes - 2021 [24] about the treatment of Hypertension in DM patients. Diabetic patients’ treatment of Hypertension should include drug classes that reduce Cardiovascular (CV) events in patients with diabetes. C09A or C09C are recommended as first-line therapy for Hypertension in individuals with diabetes and Coronary Artery Disease (CAD). Therefore, the mentioned co-prescription is aligned with diabetes treatment guidelines.

**Fourth level of ATC**

First, as shown in the results, there is a strong AR indicating the prescription of HMG CoA Reductase Inhibitors (C10AA) with Drugs Used in Diabetes. According to the Standards of Medical Care in Diabetes - 2021 [24], T2DM patients have an increased prevalence of Lipid Abnormalities, contributing to their high risk of Atherosclerotic Cardiovascular Disease (ASCVD). Statin therapy has proven effects on ASCVD consequences in subjects with and without Congenital Heart Disease (CHD). Statins are the drugs of choice for Low-Density Lipoprotein (LDL) Cholesterol-lowering and Cardio-protection for both primary and secondary prevention. Therefore, the above-mentioned AR suggests that physicians adhere to the treatment guidelines.

Second, according to our results, the prescription of Platelet Aggregation Inhibitors excl. Heparin (B01AC) is meaningfully associated with Drugs Used in Diabetes. In fact, it is proved that diabetic patients are more likely to develop Coronary and Peripheral Vascular diseases than non-diabetic subjects, and B01AC is known for the primary and secondary prevention of vascular events [18]. That is the reason why the prescription of B01AC is common in these patients. Therefore, our result is aligned with the diabetes treatment approach.

Third, the results of the ARM analysis of the fourth level of ATC showed that the prescription of Organic Nitrates (C01DA) is highly associated with Drugs Used in Diabetes. Since diabetes-associated vascular dysfunction, as well as Nitrate Tolerance, adequately responds to antioxidant therapy, this may indicate the efficacy of prescribing Organic Nitrates in diabetic patients [25].

#### **Fifth level of ATC**

First, among Angiotensin II Receptor Blockers (ARBs), Plain (C09C) drugs, Losartan (C09CA01) as an RHS had the most significant lift. Although the main effects of Losartan are due to its ATC class and are the same as those of other approved ARBs, it also has some unique benefits. For instance, a shorter duration of action, uricosuric effect, attenuation of platelet aggregation, and protective effect on the kidney [26]. Our results suggest that physicians preferred prescribing an ARB with some effects on other diabetes complications.

Second, Hydrochlorothiazide (C03AA03) and Furosemide (C03CA01) are not commonly prescribed for controlling high blood pressure. However, the results show strong lift values for these drugs. It is mainly due to their Renoprotective [27] and Cardioprotective [28] effects, which are needed to treat common diabetes complications such as Renal Insufficiency and Heart Failure.

Third, the extracted ARs of the fifth level of ATC indicate that Thiamine (vit B1) (A11DA01) is more likely to be prescribed to diabetic than non-diabetic patients. The clinical effectiveness of Thiamine (vit B1) in controlling diabetic complications has been investigated in several studies [29, 30]. According to their results, this vitamin has an effective role in Diabetic Endothelial Vascular Diseases, Lipid Profile, Nephropathy, Cardiopathy, Retinopathy, and Neuropathy.

Fourth, the findings of co-prescription patterns in the fifth level of ATC have evidenced the strong bond between Gabapentin (N03AX12) and Drugs used in diabetes. Although Gabapentin is primarily used as an anti-epileptic agent, there is an increasing pattern of its utilization as an initial pharmacologic treatment in diabetic neuropathy [31]. The added value of the hierarchical analysis of co-prescription patterns performed in this study was finding such strong rules hidden throughout previous steps.

Fifth, another interesting finding is that Thiamine (vit B1) and Gabapentin are usually associated with the prescription of Insulin which is a sign of more advanced stages of diabetes.

#### **Age- and gender-wise findings**

The most noticeable difference found in the gender-wise analysis was the higher lift value of Thyroid Preparations in females compared to males. Although this finding is sporadically mentioned in the literature, the necessity of thyroid function screening is not well established in the clinical guideline of diabetes management [24, 32, 33].

Furthermore, our results revealed that Acetylsalicylic Acid (Aspirin) was found to be associated with some of the Drugs Used in Diabetes in the young age category of males, whereas this association was not observed in young females. This gender difference was also evident in the middle-aged group, where the number of extracted ARs for females was almost half that of males, and the average support, confidence, and lift for

middle-aged males were higher than those for females. However, in the old age category, the number of rules for females increased and even slightly exceeded that for males, and we observed slight differences in quality measurements of the mentioned AR for males and females. This may suggest that the gender differences in the association between Aspirin and Drugs Used in Diabetes become less pronounced as individuals age.

In age-wise analysis, the results revealed that as one gets older, Vit B1 appears in the top rules, which may be an indication of neuropathy onset.

### **Overall findings and possible implications**

Through data mining, we revealed the current pattern of diabetes management in a developing country, and showed that the care provided in Iran is relatively evidence-based and physicians mostly adhere to guidelines. This is an important finding, as it suggests that the healthcare system in Iran is providing appropriate diabetes care despite limited resources. Additionally, we identified the order of most frequent comorbidities associated with diabetes, which can inform care by providing healthcare decision-makers and strategists with valuable insights to prioritize their efforts and resources towards the most prevalent comorbidities. This may lead to more effective management of diabetes. Furthermore, based on our findings, we would like to suggest that diabetes guidelines consider including thyroid function screening, particularly for females with diabetes. It seems to be crucial for physicians to consider lab tests of thyroid disorders in this population, as this could have significant implications for preventive care and diabetes management.

### **Strengths and limitations**

The results are based on data from the total rather than a sample of insured patients in Isfahan city; therefore, the risk of selection bias is low. To the best of our knowledge, previous similar studies were conducted with roughly a thousand samples [34–39]. In this study, we examined the prescriptions of 91,505 diabetic patients, which is way beyond the sample size of other research projects in this discipline. Additionally, the mentioned studies only found the prescription patterns regarding one level of ATC. However, we drilled down into all ATC levels for a comprehensive analysis.

Since our database is comprised of claims data, all patients were insured, and the results might not apply to the general population. In addition, some insured patients would have preferred to purchase their medications over-the-counter (OTC), and therefore no medical prescription was captured for them. Moreover, the lack of ICD-Code was one of the main limitations of this study, and this challenge was met by implementing a rule framework based on medical experts' opinions.

### **Conclusions**

The analysis carried out in this paper showed (I) the potential of association rules for pattern discovery and mining of healthcare databases because the majority of the ARs were clinically meaningful, (II) physicians' drugs of choice were mostly aligned with the Standards of Medical Care in Diabetes - 2021 [24], (III) the capability of insurance claims database as a proxy for clinical diagnoses, (IV) the effectiveness of implementing level-based ARM to find meaningful rules which were difficult to spot.

### Abbreviations

AR	Association Rule
ARM	Association Rule Mining
ASCVD	Atherosclerotic Cardiovascular Disease
ATC	Anatomical Therapeutic Chemical
CAD	Coronary Artery Disease
CHD	Congenital Heart Disease
CRISP-DM	Cross-Industry Standard Process for Data Mining
DM	Diabetes Mellitus
CV	Cardiovascular
ICD	International Classification of Disease
IHIO	Iran Health Insurance Organization
LDL	Low-Density Lipoprotein
LHS	Left-Hand-Side
MCAR	Missing Completely At Random
OADs	Oral Anti Diabetic agents
OTC	Over-The-Counter
PCOS	Polycystic Ovary Syndrome
RHS	Right-Hand-Side
STEPS	STEPwise Approach to NCD Risk Factor Surveillance
T1DM	Type 1 Diabetes Mellitus
T2DM	Type 2 Diabetes Mellitus

### Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13040-023-00344-y>.

#### Additional file 1.

### Acknowledgements

We would like to express our gratitude to the IHIO for providing us with the data used in this research. We are also grateful to Iran's Non-Communicable Diseases Research Center (NCDRC) for acting as liaison to IHIO.

### Authors' contributions

Conceptualization: EZ, SS, FF, and M-RM. Formal Analysis, Methodology, Investigation, and Software: EZ and M-RM. Funding acquisition: there is no funding. Project administration and Supervision: SS and FF. Validation: M-RM and FF. Visualization and Writing – original draft: EZ. Writing – review & editing: SS, FF, and M-RM. All authors read and approved the manuscript.

### Funding

This study was not financially funded by any organization.

### Availability of data and materials

The data used in this study is owned by IHIO. Therefore, authors are not allowed to share the data publicly or privately. However, any researcher with written permission from IHIO can request to obtain the anonymized data.

### Declarations

#### Ethics approval and consent to participate

We analyzed a de-identified secondary registry data, and no human subject was involved in this study directly. Therefore, ethics approval and consent for participation is not applicable for this study.

#### Consent for publication

All authors have consent for publication.

#### Competing interests

The authors declare no competing interests.

Received: 27 February 2023 Accepted: 20 September 2023

Published online: 20 October 2023

### References

1. Global report on diabetes. Available from: <https://www.who.int/publications/i/item/9789241565257>. Cited 2022 Nov 23.
2. IDF Atlas 10th Edition 2021.pdf. Available from: [https://diabetesatlas.org/idfawp/resource-files/2021/07/IDF\\_Atlas\\_10th\\_Edition\\_2021.pdf](https://diabetesatlas.org/idfawp/resource-files/2021/07/IDF_Atlas_10th_Edition_2021.pdf). Cited 2022 Nov 23.
3. Steps Forest - States. Available from: <https://vizit.report/panel/steps/en/main.html#/forestSex>. Cited 2022 Nov 23.
4. Lovic D, Piperidou A, Zografou I, Grassos H, Pittaras A, Manolis A. The growing epidemic of diabetes mellitus. *Curr Vasc Pharmacol*. 2020;18(2):104–9.

5. Worldwide burden of diabetes - PMC. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4192997/>. Cited 2022 Nov 23.
6. Hand DJ. Principles of data mining. *Drug Saf.* 2007;30(7):621–2.
7. Khan FA, Zeb K, Al-Rakhani M, Derhab A, Bukhari SAC. Detection and prediction of diabetes using data mining: a comprehensive review. *IEEE Access.* 2021;9:43711–35.
8. Park SH, Lee JY, Kim S. A methodology for multivariate phenotype-based genome-wide association studies to mine pleiotropic genes. *BMC Syst Biol.* 2011;5(2):S13.
9. Wright AP, Wright AT, McCoy AB, Sittig DF. The use of sequential pattern mining to predict next prescribed medications. *J Biomed Inform.* 2015;1(53):73–80.
10. Sacchi L, Dagliati A, Segagni D, Leporati P, Chiovato L, Bellazzi R. Improving risk-stratification of Diabetes complications using temporal data mining. In: 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). 2015. p. 2131–4.
11. Bhardwaj R, Datta D. Development of a recommender system HealthMudra using blockchain for prevention of diabetes. In: Mohanty SN, Chatterjee JM, Jain S, Elngar AA, Gupta P, editors. Recommender system with machine learning and artificial intelligence. 1st ed. Wiley; 2020. p. 313–27. Available from: <https://onlinelibrary.wiley.com/doi/10.1002/9781119711582.ch16>. Cited 2022 Nov 23.
12. Liu H, Xie G, Mei J, Shen W, Sun W, Li X. An efficacy driven approach for medication recommendation in type 2 diabetes treatment using data mining techniques. *Stud Health Technol Inform.* 2013;1(192):1071.
13. Lee BJ, Kim JY. Identification of type 2 diabetes risk factors using phenotypes consisting of anthropometry and triglycerides based on machine learning. *IEEE J Biomed Health Inform.* 2016;20(1):39–46.
14. Kim HS, Shin AM, Kim MK, Kim YN. Comorbidity study on type 2 diabetes mellitus using data mining. *Korean J Intern Med.* 2012;27(2):197–202.
15. Deja R, Froelich W, Deja G. Differential sequential patterns supporting insulin therapy of new-onset type 1 diabetes. *Biomed Eng Online.* 2015;14(1):13.
16. Multimorbidity—a defining challenge for health systems - The Lancet Public Health. Available from: [https://www.thelancet.com/journals/lanpub/article/PIIS2468-2667\(19\)30222-1/fulltext](https://www.thelancet.com/journals/lanpub/article/PIIS2468-2667(19)30222-1/fulltext). Cited 2022 Nov 23.
17. Masnoon N, Shakib S, Kalisch-Ellett L, Caughey GE. What is polypharmacy? A systematic review of definitions. *BMC Geriatr.* 2017;17(1):230.
18. Peron EP, Ogbonna KC, Donohoe KL. Antidiabetic medications and polypharmacy. *Clin Geriatr Med.* 2015;31(1):17–27.
19. Wirth R, Hipp J. CRISP-DM. Towards a standard process model for data mining. In: Proceedings of the 4th international conference on the practical applications of knowledge discovery and data mining (Vol. 1). 2000. p. 29–39.
20. Nahler G. anatomical therapeutic chemical classification system (ATC). In: Nahler G, editor. Dictionary of pharmaceutical medicine. Vienna: Springer; 2009. p. 8–8. [https://doi.org/10.1007/978-3-211-89836-9\\_64](https://doi.org/10.1007/978-3-211-89836-9_64). Cited 2022 Nov 23.
21. Mining frequent patterns without candidate generation | ACM SIGMOD Record. Available from: <https://dl.acm.org/doi/abs/10.1145/335191.335372>. Cited 2022 Nov 23.
22. Malekpour MR, Abbasi-Kangevari M, Shojaaee A, Moghaddam S, Ghamari SH, Rashidi MM, et al. Effect of the chronic medication use on outcome measures of hospitalized COVID-19 patients: evidence from big data. *Front Public Health.* 2023;11. Available from: <https://www.frontiersin.org/articles/10.3389/fpubh.2023.1061307>. Cited 2023 Feb 18.
23. Apache Spark™ - Unified Engine for large-scale data analytics. Available from: <https://spark.apache.org/>. Cited 2022 Nov 24.
24. Standards of Medical Care in Diabetes—2021 Abridged for Primary Care Providers | Clinical Diabetes | American Diabetes Association. Available from: <https://diabetesjournals.org/clinical/article/39/1/14/32040/Standards-of-Medical-Care-in-Diabetes-2021>. Cited 2022 Nov 23.
25. Organic nitrates and nitrate resistance in diabetes: the role of vascular dysfunction and oxidative stress with emphasis on antioxidant properties of pentaerythrityl tetranitrate. Available from: <https://www.hindawi.com/journals/jdr/2010/213176/>. Cited 2022 Nov 23.
26. Fifteen years of losartan: what have we learned about losartan that can benefit chronic kidney disease patients? - PMC. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3108782/#b3-ijnrd-3-093>. Cited 2022 Nov 23.
27. Renoprotective effects of thiazides combined with loop diuretics in patients with type 2 diabetic kidney disease - PubMed. Available from: <https://pubmed.ncbi.nlm.nih.gov/24821289/>. Cited 2022 Nov 23.
28. Cunha FM, Pereira J, Marques P, Ribeiro A, Bettencourt P, Lourenço P. Diabetic patients need higher furosemide doses: a report on acute and chronic heart failure patients. *J Cardiovasc Med.* 2020;21(1):21–6.
29. vinhquocLuong K, Nguyen LTH. The impact of thiamine treatment in the diabetes mellitus. *J Clin Med Res.* 2012;4(3):153–60.
30. Thornalley PJ. The potential role of thiamine (vitamin B1) in diabetic complications. *Curr Diabetes Rev.* 2005;1(3):287–98.
31. Bennett MI, Simpson KH. Gabapentin in the treatment of neuropathic pain. *Palliat Med.* 2004;18(1):5–11.
32. Jali MV, Kamar S, Jali SM, Pawar N, Nalawade P. Prevalence of thyroid dysfunction among type 2 diabetes mellitus patients. *Diabetes Metab Syndr Clin Res Rev.* 2017;1(11):S105–8.
33. Khassawneh AH, Al-Mistarehi AH, ZeinAlaabdin AM, Khasawneh L, AlQuran TM, Kheirallah KA, et al. Prevalence and predictors of thyroid dysfunction among type 2 diabetic patients: a case-control study. *Int J Gen Med.* 2020;12(13):803–16.
34. Palaian S, Shankar PR, Mishra P. Prescribing pattern in diabetic outpatients in a tertiary care teaching hospital in Nepal. *J Clin Diagn Res.* 2006;30:3.
35. Shamna M, Karthikeyan M. Prescription pattern of antidiabetic drugs in the outpatient departments of hospitals in Malappuram District, Kerala. *J Basic Clin Physiol Pharmacol.* 2011;22(4). Available from: <https://www.degruyter.com/document/doi/10.1515/JBCPP.2011.021/html>. Cited 2022 Nov 23.
36. Singla R, Bindra J, Singla A, Gupta Y, Kalra S. Drug prescription patterns and cost analysis of diabetes therapy in India: audit of an endocrine practice. *Indian J Endocrinol Metab.* 2019;23(1):40–5.

37. Yusefzadeh G, Sepehri G, Goodarzi H, Shokoohi M. Prescription pattern study in type 2 diabetes mellitus in diabetic out patients in private clinics in Kerman, Iran. *Br J Med Med Res.* 2014;10(4):5144–53.
38. Ashok P, Subrahmanian VT, Raj R, Babu RR, Ramshad TP, Kevin L. Prescription pattern analysis of type II diabetes mellitus inpatients and associated co-morbidities. *J Drug Deliv Ther.* 2020;10(3):42–7.
39. Concaro S, Sacchi L, Cerra C, Fratino P, Bellazzi R. Mining healthcare data with temporal association rules: improvements and assessment for a practical use. In: Combi C, Shahar Y, Abu-Hanna A, editors. *Artificial intelligence in medicine.* Berlin, Heidelberg: Springer Berlin Heidelberg; 2009. p. 16–25. (Lecture notes in Computer Science; vol. 5651). Available from: [http://link.springer.com/10.1007/978-3-642-02976-9\\_3](http://link.springer.com/10.1007/978-3-642-02976-9_3). Cited 2022 Nov 23.

### **Publisher's Note**

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

