

### 3 Results

#### 3.1 Converting real life clinical data into analyzable format

##### 3.1.1 Details of the database

The database has approximately 200 datasets (Figure 2-3). They cover various components of hospital's day-to-day functions right from operational data to the patient level clinical information. High level of classification of data types is as below:

1. Operational datasets:
  - a. Hospital charges – In Patient (IP), Out Patient (OP)
  - b. Operation theater charges
  - c. Inventory of equipment
  - d. Doctor charges
2. Reference dictionaries
  - a. Disease codes
  - b. Ayurvedic services
  - c. Medication names
  - d. Master list of Laboratory tests
  - e. Names of city, state, countries
3. Doctor details
  - a. Doctor ID
  - b. Relevant ward information
  - c. Internal / Visiting / Part time / Full time
4. Patient information
  - a. Patient details
  - b. Visit details
  - c. Vital signs
  - d. Registration details
  - e. Discharge details
  - f. Lab data details
  - g. Diet details
5. Datasets related to managing access levels, datasets related to pharmacy stocks, datasets related to scheduling appoints, datasets related to purchase orders, and other IT related contexts used by various teams in hospital

For this study, the following data was not used to in accordance with the patient data protection and privacy, financial privacy as well as hospital management confidentiality thus avoiding any controversies:

1. Hospital monetary details
2. Doctor's details (Name and ID of individual doctor)

3. Patient details of sensitive nature such as name, phone number, socio economic status, health insurance details

### 3.1.2 Data Extracted from Hospital Database

In our study, we had different versions of data, details in the table below.

Table 3-1: Versions of data used for analysis

Data version	Version 1	Version 2
Approach	CSV files provided by the Hospital IT support	Data extraction via the SQL DB connect
Date time frame	From start of the hospital to Oct 2016	From start of the hospital to Oct 2017
Data domains	Lab Vital signs Diagnosis	All the available data in the hospital database
Type of extraction	Full extraction of available domains	Full extraction of all the available hospital data

The analysis was carried out in the study using these 2 different versions of the data, version of the data has been provided along with each of the analysis for clarity.

## 3.2 Clinical data understanding

### 3.2.1 Broad checks on the datasets

This paragraph summarizes observations from structural review of the datasets. In a well-defined database, patients should have the primary key as Patient ID: mr\_no (in our case), but the underlying database considers unique visit for each patient as a primary key between tables (Patient\_ID). In general, a variable containing same information across tables should have the same name, but in our case, each table has a different variable, making it difficult to create logical links across tables. E.g., Consultation\_ID from doctor\_consultation and Patient\_ID from patient\_registration had the same information; Visit\_ID from mrd\_diagnosis and Patient\_ID from doctor\_consultation meant the same. The case report form allowed for multiple diseases and multiple treatments to be recorded for each patient, this causes a “clinical logic” challenge – the potential 1-1 relation between a disease and a treatment is lost, this had to be derived outside of the database using expert understanding which would require investment of time and efforts from Ayurvedic vaidyas. There were multiple versions of the same table available in the database (as a programmer, it is well understood that older copies are retained in the system), but

due to unavailability of the documentations increased the complexity. Potential approaches to address these challenges are provided in chapter 4 section [4.2.8](#).

This section outlines observations from the clinical data review of individual case report forms:

Vital sign dataset: Vital sign measurements include parameters like body temperature, blood pressure, body surface area, height, and weight. The existing database has various vital signs parameters listed one below the other. The current structure has one record per patient per visit per parameter. For a lot of visits vital sign information was missing, or partially filled. There were certain records with implausible values for certain parameters such as height and weight having 0 value. Blood pressure values having character data. Potential approaches to address these challenges are provided in chapter 4 sections [4.2.1](#), [4.2.5](#).

Lab measurement dataset: Findings were similar to the Vital signs database. Along with the patient identifier information, only laboratory test name and laboratory measurements were present. In case a patient had the laboratory investigations outside of the hospital that data got stored in a scanned image format. Apart from this the dataset did not contain the date of sample, reference ranges, laboratory parameter units, fasting status etc. A single lab test had multiple names. E.g., Alanine Aminotransferase was captured in the dataset in the following different ways:

Alanine Aminotransferase

Alanine Aminotransferase      ALT (SGPT)(UV Kinetic)

Alanine Aminotransferase (SGPT)(UV Kinetic)

Alanine Aminotransferase ALT (SGPT)

Alanine Aminotransferase ALT (SGPT)(UV Kinetic)

S.G.PT ( UV kinetic)

SGPT ( UV kinetic)

ERYTHROCYTE SEDIMENTATION RATE was captured in the dataset in the following different ways:

ERYTHROCYTE SEDIMENTATION RATE

ERYTHROCYTE SEDIMENTATION RATE ( ESR)

ERYTHROCYTE SEDIMENTATION RATE ( ESR)

Potential approaches to address these challenges are provided in chapter 4 sections [4.2.3](#), [4.2.5](#).

Treatment dataset: the treatment or dosing or medication dataset does not get exported into a structured file for easy understanding and analysis. Which treatment was prescribed for which disease was not easily understandable based on the system generated report. Potential approaches to address these challenges are provided in chapter 4 sections [4.2.4](#), [4.2.8](#).

Medical coding and clinically important variables: the medical records for patients were captured differently by different doctors, nurses and other medical staff. Same information was found in more than one variable. Acronyms were used inconsistently. Answer for more than one question was captured in one variable. Due to “free text nature” of variables simple questions like Yes / No had many different data values. Potential approaches to address these challenges are provided in chapter 4 section 4.2.5.

Classification and Sub-classification of the Doshas / Diseases: It was observed that the main disease classification by kapha, pitta, vata has disparity in numbers. As an example the table below show the variation in the counts of the diseases and their sub-classification. Potential approaches to address these challenges are provided in chapter 4 section 4.2.6

Figure 3-1: A snippet of disease table by gender

DisType_Diseases					
Distype	Code	Description	Patient Gender		
			F	M	Grand Total
Metabolic	P5.0	Prameha	648	849	1,497
	P5.3	Prameha - Sthula	23	20	43
	P5.2	Prameha - Pidaka	5	12	17
	P5.1	Prameha - Krusha	4	5	9
Grand Total			671	876	1,547

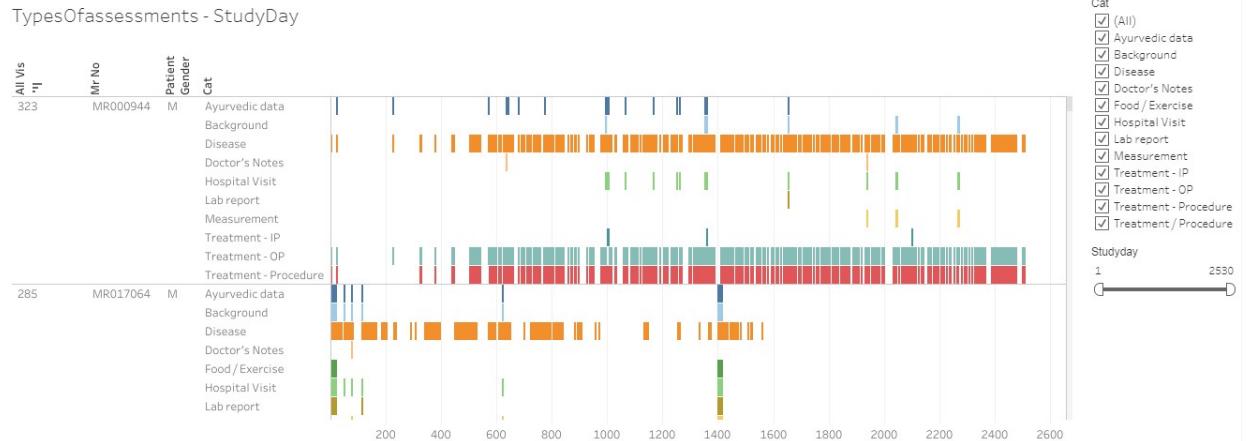
DisType_Diseases					
Distype	Code	Description	Patient Gender		
			F	M	Grand Total
RMSD	A2.0	Aamavaata	652	180	832
	A2.3	Aamavaata - Vaataja	27	7	34
	A2.1	Aamavaata - Kaphaja	13	4	17
	A2.2	Aamavaata - Pittaja	4	3	7
Grand Total			679	190	869

Frequency of Prameha (P5.0) and Aamavaata by gender has been displayed. The frequency of patients is substantially lower for classification by Sthula, Pidaka and Krusha as well as Kapha, Vaata, and Pitta. Data version: 2011 to Oct 2017. Link to analysis: [Link](#)

### 3.2.2 Contents checks

The analysis below shows that for majority of the patients and for majority of the visits, the disease data and medication (Treatment /Procedure) were non-missing. Most of the other categories were not entered as consistently as they should have been. If this was the expected data collection pattern then these findings should not be considered as any issues.

Figure 3-2: Variable classification by categories



500+ variables are captured for each visit and each patient are classified into the following categories, (1) Ayurvedic data, (2) Background, (3) Disease, (4) Doctor's Notes, (5) Food / Exercise, (6) Hospital Visit, (7) Lab report, (8) Measurement, and (9) Treatment / Procedure. If there is any non-missing data present in a particular category then “Yes” is assigned, if the data is missing then “No” value is assigned. This data is presented as a listing for each patient for each visit (day). When the data is available it is presented as a color-coded bar and when it is missing then it is presented as a white blank space. Data version: 2011 to Oct 2017. Link to analysis: [Link](#)

Two example screenshots are shown below as to what was observed while content check was performed.

CRFname_variable number_variable label	Unique values	Unique patients	Variable classification category
sec001_var008_Diabetes	1064	4124	Background

sec001\_var008\_Diabetes: represents CRF page number 1, labelled as “History of Present Illness” and variable number 8 “Diabetes”. Based on the label of the variable, this variable is considered as a part of “Background” information. This variable “Diabetes” has 1064 unique values entered by different doctors for different patients. Data is available for 4124 distinct patients.

CRFname_variable number_variable label	Unique values	Unique patients	Variable classification category
sec001_var018_Associated Complaint with Onset & Duration	5102	4549	Disease

sec001\_var018\_Associated Complaint with Onset & Duration: presents CRF page 1 “History of Present Illness” and variable number 8 “Associated Complaint with Onset & Duration”. This variable “Associated Complaint with Onset & Duration” has 5102 different values entered by different doctors for different patients. Data is available for 4549 distinct patients.

The large number of unique values show that the data entry rules are not being followed consistently, and each doctor or each nurse might have a different interpretation of the rules. In addition to this, looking at the unique number of patients, the data has not been entered for all the patients, hypothetically giving rise to missing data. Potential approaches to address these challenges are provided in chapter 4 section 4.2.8.

### 3.2.3 Visit pattern analysis

Each of the cell displayed on a calendar display was coloured in shades of blue from light blue to dark blue showing increasing frequency count of number of patients. From 2011 to 2016, the number of patients visiting hospital on weekdays was less than the number of patients visiting on weekends. In-Patients were considerably less than Out-Patients. Overall number of patients coming to hospital have been increasing year on year. Potential approaches to address additional patient burden are provided in chapter 4 section 4.2.1.

Figure 3-3: Visit pattern analysis

Sheet 1



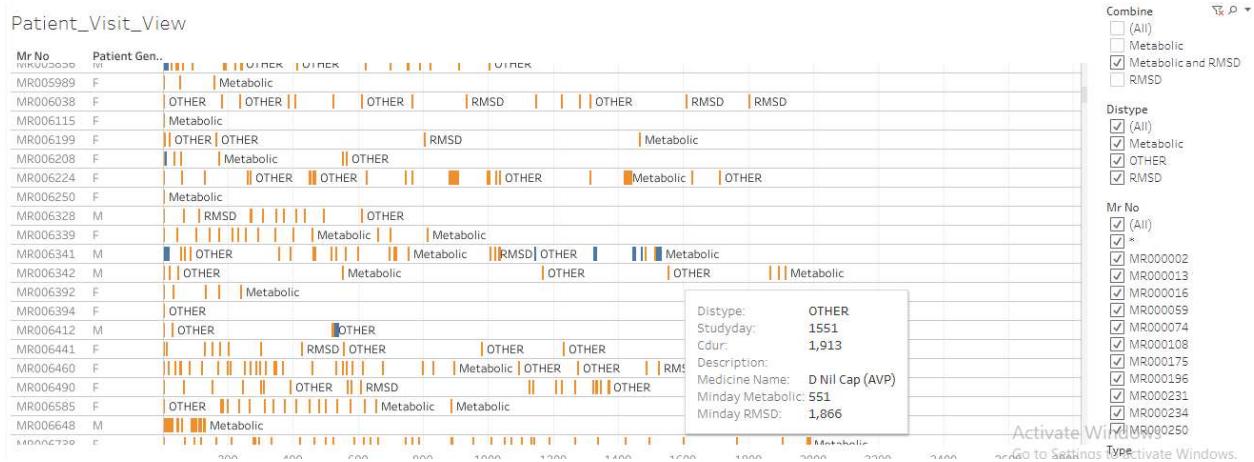
Frequency counts of 4 parameters, (1) new Out-Patients added on that day, (2) total number of patients visiting on that day, (3) total number of In-Patient visits on that day, and (4) total number of Out-Patient visits on that day are calculated for each day to understand the patient flow to hospital from year on year. Light blue to dark blue shows increasing frequency count of number of patients Data version: 2011 to Oct 2016. Link to analysis: [Link](#)

### 3.2.4 Patient disease and treatment journey view

Multiple representations of data allow the end user to review data with different perspectives. Patient profile reports provide detailed view of individual patient's disease condition, prescribed medication, co-morbidities along with basic demographic information. Treating doctors and researchers will greatly benefit from this visual display. This representation (Figure 3-4) provides the patient an understanding of the disease chronology as well as the prescribed

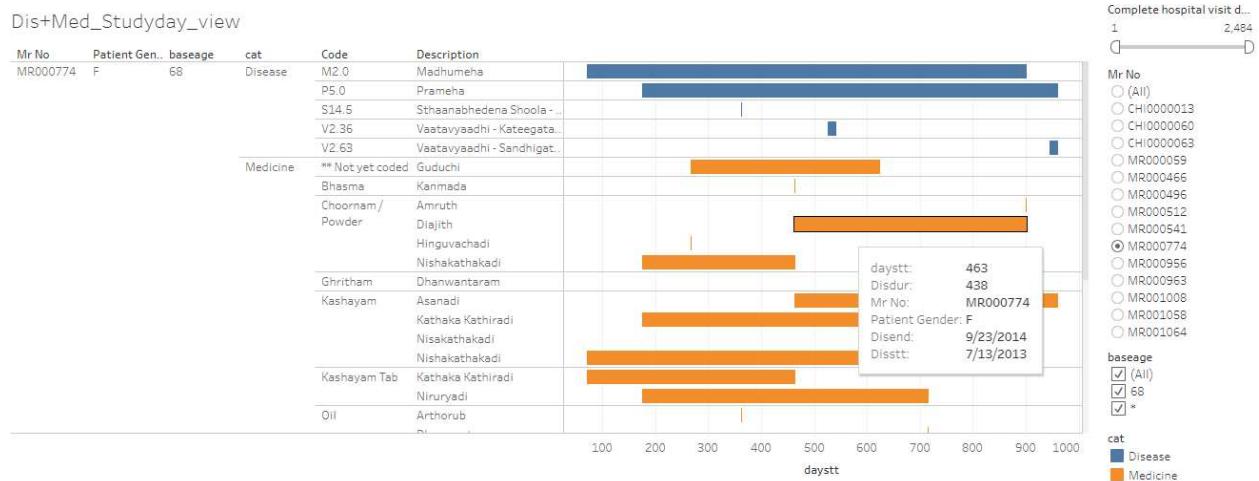
medication and progress. Interpretations drawn from these representations and potential approaches to improve current version of patient profiles used at the hospital are provided in chapter 4 sections 4.2.8.

Figure 3-4: Patient visit profile – Horizontal view



Mr No: Patient ID, Patient gender, x-axis: duration of hospital visits, Orange bar: Out-patient visit, Blue bar: In-patient visit, Metabolic: when a metabolic disease is reported, RMSD: when a Rheumatic and Musculoskeletal disease is reported, OTHER: other diseases are reported. Patients listed have at least one of the RMSD or metabolic diseases. Tooltip has a lot of information relevant to each visit. Data version: 2011 to Oct 2017. Link to analysis: [Link](#)

Figure 3-5: Patient visit profile – Vertical view



Mr No: Patient ID, Patient gender, baseage Age at the very first hospital visit, category: Disease and medicine, Code: ACD code, Description: disease description, x-axis: duration of disease and medicine, Tooltip has a lot of information relevant to each visit. Data version: 2011 to Oct 2017. Link to analysis: [Link](#) to get the above display subset Mr No = MR000774

### 3.3 Studying demographics and patient specific factors

Results related to complete set of patients: While exploring the basic data, the following high-level picture appeared: For the 5-year time frame from 2011 to 2016, the database contained

approximately 40,000 unique patients (Figure 3-6), 90% of patients were from India and remaining 10% patients were from more than 50 different countries (Figure 3-7). The proportion of male and female patient was approximately 50%. Median age for females was marginally higher than males across all visit types (Figure 3-8). Approximately 90% of patients were Out-Patients and 10% were In-Patients (Figure 3-1). Approximately 12,000+ female patients and 14,000+ male patients had reported only a single disease (Figure 3-11), these patients could have come only once to the hospital and may not have come back at all after reporting the first disease. There were a few outliers observed having more than 10 disease conditions across the years. The maximum age of 108 years was a possible case of data issue. Similar anomalies were seen in a few other groups, e.g., patients reporting 23 diseases, is this accurate? This warrants additional data checks from operational and clinical perspective. Blood group is collected only for ~32,500 out of ~40,000 patients. There was missing data for almost 20% of patients. Blood group distribution was largely in line with the Indian blood group distribution (Figure 3-9). Potential approaches to improve current situation and interpretations are provided in chapter 4 section 4.3.

Figure 3-6: Total Number of Patients

01 NoOfPatients	02 Country	03 AgeBoxplotCountry	04 Age
-----------------	------------	----------------------	--------

## Summary of distinct number of patients

Grp				
Common	NA	OnlyIP	OnlyOP	Grand Total
3,491	0	4,145	38,903	39,557

Grad Total: Total number of patients used in the analysis, OnlyIP: patients having only In-Patient visits, OnlyOP: patients having only Out-Patient visits, Common: Patients having both type of visits. Data version: 2011 to Oct 2016. Link to analysis: [Link](#)

Figure 3-7: Country-wise Visualization

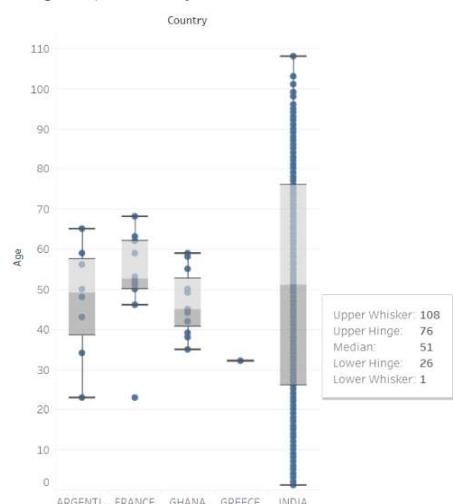
## 02 Country



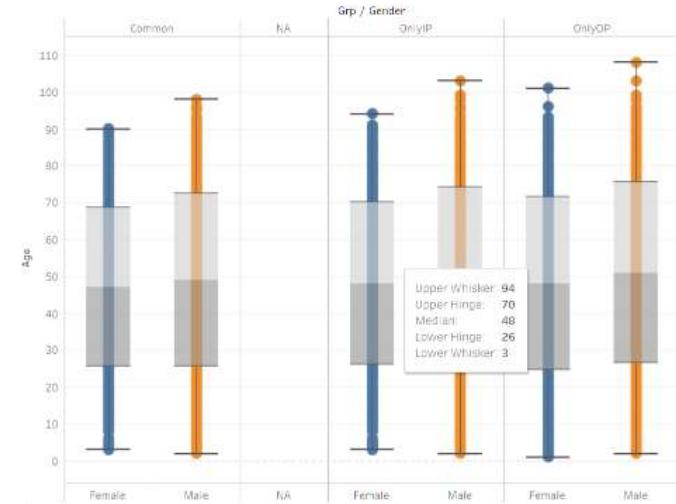
Unique number of patients are plotted on the world map, the map shows that at least 1 patient data is coming from 50+ countries, 95% or more patients are from India. Data version: 2011 to Oct 2016. Link to analysis: [Link](#)

Figure 3-8: Age distribution by country, age distribution by gender

## 03 AgeBoxplotCountry



## 04 AgeBoxplotGroup



Boxplot representation: Age distribution is presented for each country and then by type of patient and gender, OnlyIP: patients having only In-Patient visits, OnlyOP: patients having only Out-Patient visits, Common: Patients having both type of visits Data version: 2011 to Oct 2016. Link to analysis: [Link01](#), [Link02](#)

Figure 3-9: Blood-group Distribution by gender

## 07 BloodGroup

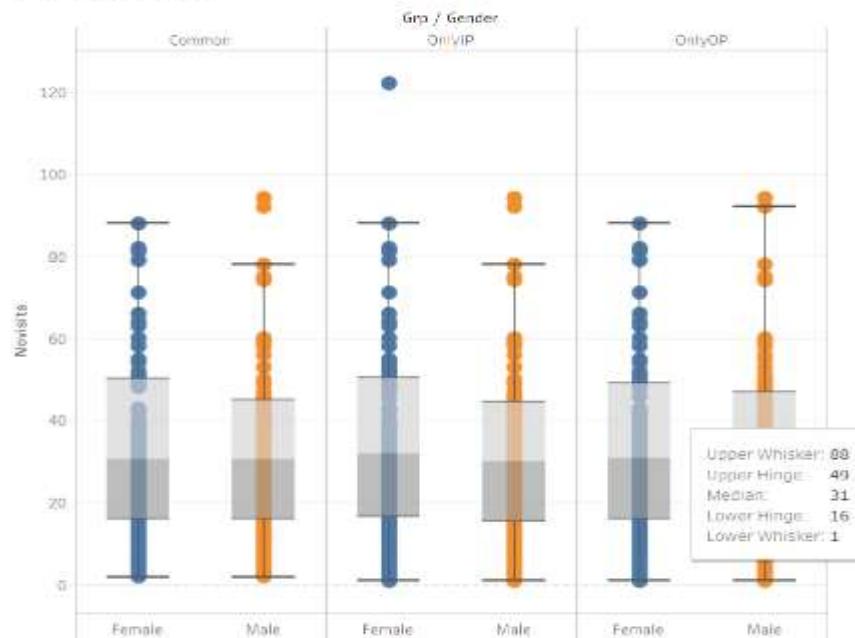
Blood... 	Gender			Grand Total
	Female	Male	NA	
O+	6,340	6,684		13,024
B+	4,490	4,956		9,446
A+	2,997	3,030		6,027
AB+	828	1,046		1,874
O-	434	447		881
B-	265	259		524
A-	239	178		417
AB-	80	87		167
A1 +ve	30	17		47
A1 -ve	6	1		7
A1B +ve	3	4		7
A2 +ve	1			1
A2B +ve		1		1
NA			1	1
Null	0	0		0
Grand Total	15,713	16,710	1	32,424

Tabular frequency distribution table for Blood-group by gender. Data version: 2011 to Oct 2016. Link to analysis: [Link](#)

Figure 3-10: Number of Visits, and Visit Types

01 NoOfPatients | 02 Country | 03 AgeBoxplotCountry | 04 AgeBoxplotGroup | 05 NoOfVisitsBox | 05a NoC

05 NoOfVisitsBox



Boxplot representation of number of Visits, and Visit Types, Data version: 2011 to Oct 2016. Link to analysis: [Link](#)

Figure 3-11: Descriptive summary statistics by number of Diseases by Age and Gender

01 NoOfPatients | 02 Country | 03 AgeBoxplotCountry | 04 AgeBoxplotGroup | 05 NoOfVisitsBox | 05a NoOfDis\_age | 05b N

05a NoOfDis\_age

Noofdiseases	Gender													
	Female			Male										
	Count	Age	Avg_Age	Median	Std.dev.	Age	Count	Age	Avg_Age	Median	Std.dev.	Age	Min_Age	Max_Age
1	12,884	46	46	46	18	1	96	45	45	45	19	2	108	
2	5,652	48	50	50	18	2	93	48	48	47	19	2	96	
3	2,426	48	48	48	17	2	93	50	49	49	18	2	98	
4	1,215	51	51	51	17	2	87	53	53	55	19	2	93	
5	526	50	50	50	15	9	84	53	53	54	18	2	87	
6	342	48	48	48	17	9	81	50	50	47	15	6	84	
7	182	53	52	52	13	12	79	57	57	56	15	33	88	
8	99	48	48	48	15	13	80	46	46	52	23	6	76	
9	77	54	52	52	15	4	78	58	58	65	11	39	67	
10	25	53	51	51	21	33	86	50	50	51	5	43	57	
11	23	47	49	49	17	22	66	44	44	44	14	35	63	
12	31	44	39	39	24	9	82	64	64	64	0	64	64	
15	2	56	56	56	21	41	71							
NA	4,263	46	45	45	19	2	101	43	43	41	19	2	103	
13	8	53	56	56	9	30	58	51	51	47	8	31	58	
23							6	41	41	41	0	41	41	
14	9	30	18	18	20	18	71	1	35	35	35	35	35	

Descriptive summary statistics by number of diseases reported, by gender. Noofdiseases: Number of diseases reported in the database. Data version: 2011 to Oct 2016. Link to analysis: [Link](#)

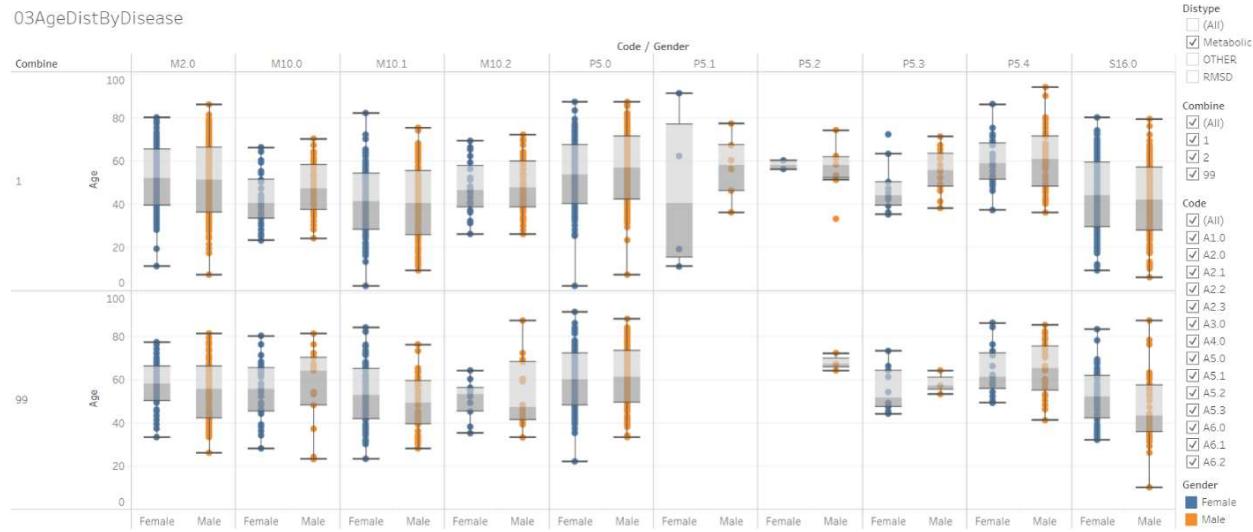
Results for the metabolic and RMSD disease areas: Out of ~40,000 patients, there were ~14,000 patients having reported at least 1 metabolic and/or 1 RMSD disease condition. It was quite evident that there were a lot more patients in the RMSD group compared to the metabolic group

(Figure 3-12). Large number of patients were visiting the hospital only for 1 visit, ~62% patients were dropping off in first month of treatment. ~15% of patients having at least one RMSD disease were still visiting the hospital after 1 year of first ever visit to the hospital. Boxplot representation of age showed variability in age across disease type and gender (Figure 3-13). Presentation of disease burden by gender, Indian seasons (ratus) and disease category provides data about possible variations reported for different diseases (Figure 3-15). (1) Prameha, (2) Madhumeha, and (3) Sthaulya were the top three most frequently reported metabolic diseases where as (1) Vaatavyaadhi – Sandhigata Vaata, (2) Vaatavyaadhi, (3) Vaatavyaadhi – Gridhrasee, (4) Sthaanabhedana Shoola – Katee Shoola and (5) Sthaanabhedana Graha – Katee Graha were the top five most frequently reported RMSD diseases. Prameha and Madhumeha were reported more by males than females. There were more female patients with disease condition Sthaulya. In general, RMSD diseases were reported in more females than males. For RMSD disease group, 51 out of 97 diseases were reported in <= 10 patients. Metabolic diseases were not varying across seasons, while RMSD diseases had some seasonal variations (Figure 3-15). The before and after visualization of data allows to build a disease and medicinal trajectories (Figure 3-16). These should be useful for determining diagnostic and prognostic relationships.

Figure 3-12: Data tabulation for patients reporting RMSD and Metabolic diseases



Figure 3-13: Disease distribution by age and gender



Boxplot representation of age by disease. Orange: male, Blue: female. Individual column represents a disease. 1 = Patients with at least 1 metabolic diseases, 2 = Patients with at least 1 Rheumatic, Musculoskeletal (RMSD) diseases, 99 = Patients with at least one disease from each of the groups, Data version: 2011 to Oct 2016. Link to analysis: [Link](#)

The columns in the above image are different diseases from Metabolic and RMSD categories. (Refer [Link](#) for codes and descriptions)

Figure 3-14: Patient visit duration for Disease categories by Gender

Durwr	Dur	Combine / Gender								Grand Total	
		1		2		99					
Female	Male	Female	Male	Female	Male	Female	Male	Female	Male		
Null	Null	42	49	275	228	53	41				688
0	>=1 day	1,134	1,537	5,751	4,763	582	447				14,214
1	>=1 month	425	590	2,255	1,634	354	267				5,525
2	>=2 months	334	480	1,854	1,374	315	239				4,596
3	>=3 months	294	433	1,650	1,227	292	225				4,121
6	>=6 months	223	322	1,257	944	249	196				3,191
12	>=1 year	148	203	857	665	184	147				2,204
24	>=2 years	81	103	444	335	102	95				1,160
36	>=3 years	38	48	215	159	55	51				566
48	>=4 years	17	16	62	47	24	19				185
60	>=5 years	4	2	18	11	10	6				51
Grand Total		1,147	1,554	5,821	4,821	587	453				14,383

1 = Patients with at least 1 metabolic diseases, 2 = Patients with at least 1 Rheumatic, Musculoskeletal (RMSD) diseases, 99 = Patients with at least one disease from each of the groups, Data version: 2011 to Oct 2016. Link to analysis: [Link](#)

Figure 3-15: Disease distribution by Seasonal Variations and gender

Distype	Code	Description	season / Patient Gender																	
			01 Vasant Rutu		02 Grishma Rutu		03 Varsha Rutu		04 Sharad Rutu		05 Hemant Rutu		06 Shishir Rutu		Grand Total					
Metabolic	P5.0	Prameha	163	207	194	263	208	279	167	230	151	187	166	203	1,497					
	M2.0	Madhumeha	158	268	184	271	188	309	198	330	144	241	153	244	1,441					
	S16.0	Sthaulya	109	59	147	105	138	94	161	95	111	81	119	75	989					
	M10.1	Medoroga - Sthula medho ..	88	53	97	52	98	54	89	46	84	47	78	39	627					
	M10.2	Medoroga - Sukshma med..	14	32	13	38	9	40	14	36	11	29	9	23	193					
	P5.4	Prameha - Upadrava	8	21	14	27	19	28	16	15	6	20	12	15	150					
	M10.0	Medoroga	16	15	18	13	23	17	16	14	18	12	19	13	149					
	P5.3	Prameha - Sthula	6	4	5	5	3	6	6	3	6	6	3	3	43					
	P5.2	Prameha - Pidaka	3	1			7		3	1	1	1			2	17				
	P5.1	Prameha - Krusha	1					1	3	1	1	1			9					
RMSD	V2.63	Vaatavyaadhi - Sandhigat..	712	317	855	384	855	376	784	334	619	284	585	276	4,299					
	V2.0	Vaatavyaadhi	373	316	415	354	485	373	368	306	275	249	362	290	2,993					
	V2.23	Vaatavyaadhi - Gridhrasee	366	308	417	362	450	361	435	335	342	293	335	275	2,613					
	S14.21	Sthaanabhedena Shoola - ..	133	133	163	139	190	154	175	154	116	114	139	116	1,333					
	S13.11	Sthaanabhedena Graha - ..	90	104	106	113	130	133	122	113	117	111	114	107	982					
	A2.0	Aamavaata	218	59	225	62	212	63	237	63	177	48	186	50	832					
	A3.0	Abhighataja Shoola	59	66	90	89	76	85	68	91	57	68	51	69	739					
	V1.0	Vaatarakta	42	58	61	83	85	75	74	73	62	63	50	47	529					
	S14.0	Sthaanabhedena Shoola	73	42	80	53	51	43	68	51	46	49	74	44	523					
	V2.36	Vaatavyaadhi - Kategata..	52	41	54	60	62	61	42	29	16	29	33	35	404					
	S14.7	Sthaanabhedena Shoola - ..	21	21	32	30	21	31	32	31	26	16	28	21	233					
	S14.17	Sthaanabhedena Shoola - ..	14	15	30	18	28	20	29	27	19	18	7	11	206					
	V2.9	Vaatavyaadhi - Asthigata ..	26	20	19	19	21	19	42	15	36	20	21	8	206					
	V2.75	Vaatavyaadhi - Vaatakant..	21	5	26	14	43	14	29	9	20	5	18	8	164					
	S13.19	Sthaanabhedena Graha - ..	14	15	12	18	18	17	12	10	12	14	15	14	121					
	V2.30	Vaatavyaadhi - Jaanugata..	12	7	20	13	10	5	7	10	4	4	8	3	92					
	S15.31	Sthaanabhedena Shoola - ..	11	9	9	7	14	7	11	16	14	6	9	5	89					
	S14.3	Sthaanabhedena Shoola - ..	4	6	15	7	0	6	12	5	0	6	6	11	62					

Distype: Metabolic and RMSD, Code: ACD disease code, Description: disease description, seasons are presented as: Vasant rutu, Grishma rutu, Varsha rutu, Sharad rutu, Hemant rutu, and Shishir rutu, Data version: 2011 to Oct 2017. Link to analysis: [Link](#)

Figure 3-16: Pre and Post Disease Classification Analysis

#### Example 1: Prameha



Cat: category of disease and medicine, Code: ACD code, prescribed medicine types, pre and post visit window w.r.to the 1<sup>st</sup> day of each of the reference diseases, Data version: 2011 to Oct 2017. Link to analysis: [Link](#) to get the display for Prameha, go to "Refcode, Refdesc" filter and select P5.0, Prameha.

#### Example 2 Vaatavyadhi – Sandhigata Vaata



Cat: category of disease and medicine, Code: ACD code, prescribed medicine types, pre and post visit window w.r.to the 1<sup>st</sup> day of each of the reference diseases, Data version: 2011 to Oct 2017. Link to analysis: [Link](#) to get the display for Vaatavyaadhi – Sandhigata Vaata, go to “Refcode, Refdesc” filter and select V2.63, Vaatavyaadhi – Sandhigata Vaata.

### 3.4 Studying diagnostics and interventions

Almost all the ICD categories were represented in the analysis. Some of the ICD classes had more patients than other categories. Age distribution showed natural variation. Visit distribution and duration for which patients were visiting hospital look like the earlier analysis. The ACD and ICD mapping exercise showed that the current hospital data demonstrates all the types of diseases being catered to at the hospital. (Figure 3-17, Figure 3-18, Figure 3-19, Figure 3-20)

The bar graph representation (Figure 3-21) provided a view of the spread of the patient population their diseases versus their prakriti classification seen at the hospital. The prakriti type as well as disease type can be filtered. It also provided a view of the combination of the gender and the prakriti manifesting into the kind of doshas and the most prevalent doshas for the combination.

Observations from the first Co-morbidity approach (Figure 3-22, Figure 3-23, Figure 3-24): This display provided a comprehensive view of the disease clusters. Comorbidities were easily identified, some of them are clinically relevant, and some of them are not. Bubble size provided comparative view of number of patients reporting a specific disease. The age group distribution for each gender was available. Some diseases were reported more by males or by females, easy to spot on the graph. Box named “Number of other diseases” provided a contextual display about number of co-morbidities. Some diseases had higher number of co-morbidities, some had lower number. Variations were seen amongst gender as well. This analysis did not consider the before or after nature of time points, hence did not provide insights into the causal relationships between diseases. The second co-morbidity approach provided views on the seasonal variations of diseases as well as seasonal co-morbidities (Figure 3-25). The third co-morbidity approach provided the following: The collapsible tree showed progression of diseases as experienced by patients and reported in the database. The tree showed approximately 12,500 lines of data in very short space. Some diseases were experienced more by males than by females. Some diseases

were only reported by one of the genders. Some diseases had many more branches than a few others. Some of the disease trajectories had very few numbers of patients. Some of the trajectories may be clinically meaningful and some may not be meaningful (Figure 3-26).

Treatment and disease analysis at individual patient level: This analysis was explained using example patient MR000335: For this patient there were 17 visits in the database, there were 17 distinct diseases reported and 45 distinct treatments, services prescribed. 4 out of 17 diseases were repeated and 10 out of 45 treatments had been repeated. When a new disease was reported, usually a new treatment or treatments re-reported, if there was only a new treatment added then it could indicate, the earlier treatment may not have worked, or it described the treatment regimen. If only new diseases were added and no new treatment was added, then the same treatment could work for multiple diseases. These visualizations allow the treating doctor insights into newer diseases getting reported as well as what newer treatments have been prescribed at what time points (Figure 3-27, Figure 3-28, Figure 3-29).

Area graph representation of diseases provided information about 800+ diseases, almost all diseases present in the database in very short space. Due to the data visualization scheme variations caused by day-to-day, seasons, gender, and diseases could be interpreted very easily. The interactive nature of visualization allowed for real time subset of diseases. One of the 4 diseases displayed has very few patients compared to other 3 diseases showing different nature of diseases (Figure 3-30).

Cross tabulation of prescribed treatments and disease group by gender was generated. First example is created using Balaristham. The source variable captured the quantity + unit + company name in the same variable, which did not allow for 100% accurate numerical calculations, but still provided a good idea. Only 30 patients having metabolic diseases were prescribed the medicine whereas 1,142 patients with RMSD were prescribed. Second example is created using bhasma: approximately 287 patients out of 17,406, 1.5% of patients are prescribed bhasmas (at least treatments having word “bhasma”) for various diseases (Figure 3-35, Figure 3-36). Summary statistics and hypothesis testing was conducted to conclude any impact of bhasma on visit duration. 514 patients were identified with at least 1 bhasma treatment. The mean duration of treatment before 1<sup>st</sup> bhasma treatment was 14.8 days, min - max duration was reported as (1, 111) whereas the mean duration of treatment after 1<sup>st</sup> bhasma treatment was 10.6 days, min - max duration was reported as (0, 89). The t-test at 5% significance level shows statistically significant difference between duration of treatment before bhasma treatment and duration of treatment after bhasma treatment.

Circular display, how to read the visualization? Figure 3-37 and Figure 3-38 show 2 examples of 2 combinations, the first example had many green bars, and the second combination had very few green bars. Details about the display: For each reference disease 1 page was created. Each page was controlled by a combination of “Reference disease + disease”, “Reference disease + medicine”

Refcode, Refdesc

V2.23, Vaatavyaadhi ...

Show History

HighlightCode  
A6.0:Amlapitta

Reference disease or medicine window:

Tables displayed in the top part of the display: there were 9 columns created for each time point.

01 Day 1 District count of Tab Code	02 1 month District count of Tab Med	03 2 months District count of Tab Code	04 3 to 6 months District count of Tab Code	05 7 to 12 months District count of Tab Code	06 2nd year District count of Tab Med	07 3rd year District count of Tab Med	08 4th year District count of Tab Code	09 5 years District count of Tab Med
---	--	--	---	--	---	---	--	--

2 columns were displayed in each time point to display “count of distinct number of diseases” and “count of distinct number of medicines”.

Count of distinct number of diseases:

Distinct count of Tab Code
Distinct count of Tab Med

Count of distinct number of medicines:

Distinct count of Tab Code
Distinct count of Tab Med

There were 3 rows for “Before period”, “Day 1” and “After period” with 2 lines in each period. Day 1 cell showed the start day of reference disease “V2.23: Vaatavyadhi – Gridhrasee”, This example showed 137 total number of distinct diseases reported and 664 total number of distinct medicines prescribed on day 1 for this combination of reference disease “V2.23: Vaatavyadhi – Gridhrasee” and disease “A6.0: Amlapitta”.

The first line in the Day 1 cell showed, 1 disease – which is “A6.0: Amlapitta” and 81 distinct medicines prescribed. These 81 different treatments could have been prescribed for “A6.0: Amlapitta”.

		01 Day 1	
		Distinct count of Tab Code	Distinct count of Tab Med
01	In		
Before	O.		
	T..		
02 Day	In	1.0	81.0
1	O.	136.0	664.0
	T..	137.0	664.0

Cells in the “Before period” line provided the following information:

This example shows 1 month before “V2.23: Vaatavyadhi – Gridhrasee”, there were 61 distinct diseases and 298 distinct medicines reported.

The first line in the cell shows, 1 disease – which is “A6.0: Amlapitta” and 11 distinct medicines prescribed. These 11 different treatments could have been prescribed for “A6.0: Amlapitta”.

		01 Day 1		02 1 month	
		Distinct count of Tab Code	Distinct count of Tab Med	Distinct count of Tab Code	Distinct count of Tab Med
01	In			1.0	11.0
Before	O.			60.0	294.0
	T..			61.0	298.0

Cells in the “After period” line provide the following information:

This example shows 1 month before “V2.23: Vaatavyadhi – Gridhrasee”, there were 96 distinct diseases and 633 distinct medicines reported.

The first line in the cell shows, 1 disease – which is “A6.0: Amlapitta” and 49 distinct medicines prescribed. These 49 different treatments could have been prescribed for “A6.0: Amlapitta”.

		01 Day 1		02 1 month	
		Distinct count of Tab Code	Distinct count of Tab Med	Distinct count of Tab Code	Distinct count of Tab Med
03 After	In			1.0	49.0
	O.			96.0	633.0
	T..			97.0	635.0
<b>Grand Total</b>		<b>137.0</b>	<b>664.0</b>	<b>114.0</b>	<b>706.0</b>

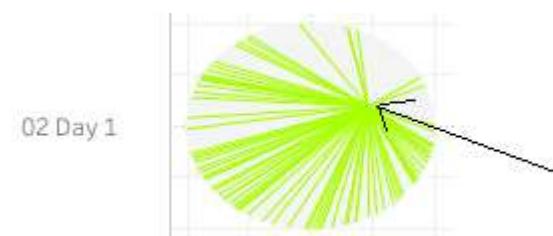
The bottom section follows the same structure as the top section. The following table provided distinct number of diseases reported and distinct number of medicines prescribed for this particular combination. There were 308 diseases and 1,357 medicines reported for this

combination of reference disease “V2.23: Vaatavyadhi – Gridhrasee” and disease “A6.0: Amlapitta”.

Distinct diseases / medicines	
Disease	Medicine
308	1,357

Explanation about the circular view:

The starting point marked the position of the other disease in this case, “A6.0: Amlapitta”, The green colored spokes going from point of origin were different treatments prescribed. These were showing 664 distinct medicines prescribed on day 1.



Hovering tooltip provided details about the disease, treatment name and count of number of



patients:

The inner circle displayed the diseases. And the outer circle displayed the treatments.

Distance based Pre and post analysis: in the example for the disease M2.0 (Madhumeha), disease trajectory distances were plotted (Figure 3-39). (1) Madhumeha was reported by 1,441 patients at least once. (2) Of these 567 patients reported only Madhumeha and no other disease thereafter. For such patients, the disease trajectory calculation was not possible, hence these patients were removed from the analysis. (3) The following table showed details of patient count. The disease trajectory calculation was based on 874 patients comprising of 530 males and 344 females

patient_gender	numtype	ncount
M	Only 1 dis	377
F	Only 1 dis	190
M	> 1 dis	530
F	> 1 dis	344

Out of 530 male patients: (1) 201 patients had diseases reported before the first reported instance of M2.0, (2) 460 patients had at least one other disease reported other than M2.0 on or after the first reported instance of M2.0, (3) Disease trajectory for 70 patients could not be calculated since the next reported disease was M2.0.

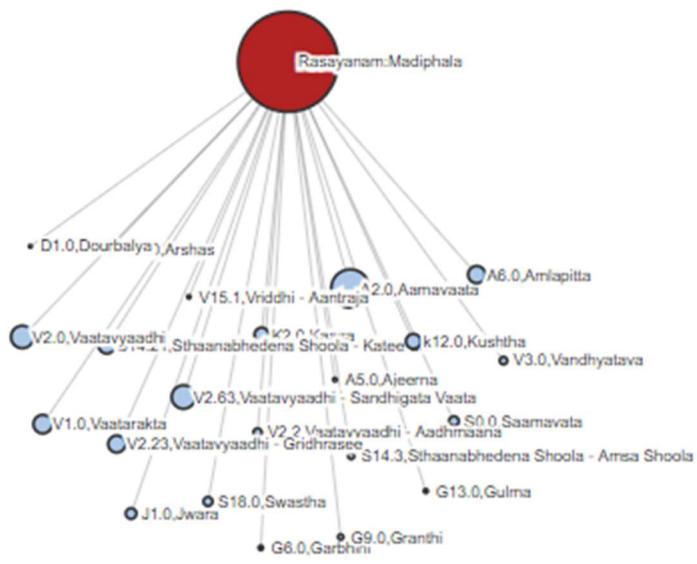
Out of 344 female patients: (1) 140 patients had diseases reported before the first reported instance of M2.0, (2) 291 patients had at least one disease other than M2.0 reported on or after the first reported instance of M2.0, (3) Disease trajectory for 53 patients could not be calculated since the next reported disease was M2.0.

patient_gender	numtype	refday2	ncount
M	> 1 dis	Before	201
M	> 1 dis	After	460
F	> 1 dis	Before	140
F	> 1 dis	After	291

The trajectories were calculated for these patients and displayed for the before and after period. More number of patients had disease trajectories in the “after onset” section. More than 73% of males lie in the score >0.25 and around 36% of them lie in the score >0.5 which could confirm that there were similar diseases experienced by the patients post the onset of the reference disease. Similarly, around 64% of females lie in the score >0.25 and around 24% of them lie in the score >0.5 which could confirm that there were similar diseases experienced by the patients post the onset of the reference disease. A few more examples of similar kind were shown for diseases: P5.0: Prameha, V2.23: Vaatavyaadhi - Gridhrasee, V2.63: Vaatavyaadhi - Sandhigata Vaata (Figure 3-40).

The radar plot shows multi-dimensional data in a short space, 7 different parameters were shown on 7 vertices. Different shapes suggest that there were underlying differences to the data structure.

Dynamic bubble plot visualization showed intricate relationships and details in a limited space, one more analysis method to see multi-dimensional data (Figure 3-42). Using data for Amavaata (A6.0), in the snapshot below, we could see the patients with Aamvata who received Rasayanam Madiphala and in addition what were the other diseases that these patients reported.



Another view which showed the further detailed view of Aamvata patients having received Rasayanam Madiphala and further having reported Stanbheedana Kathee Shoola disease and further its treatment. Further, the big bubble here closer to Stanbheedana Kathee Shoola disease was Aamdosha which showed that Stanbheedana Kathee Shoola was also one of the diseases reported by Aamdosha patients.

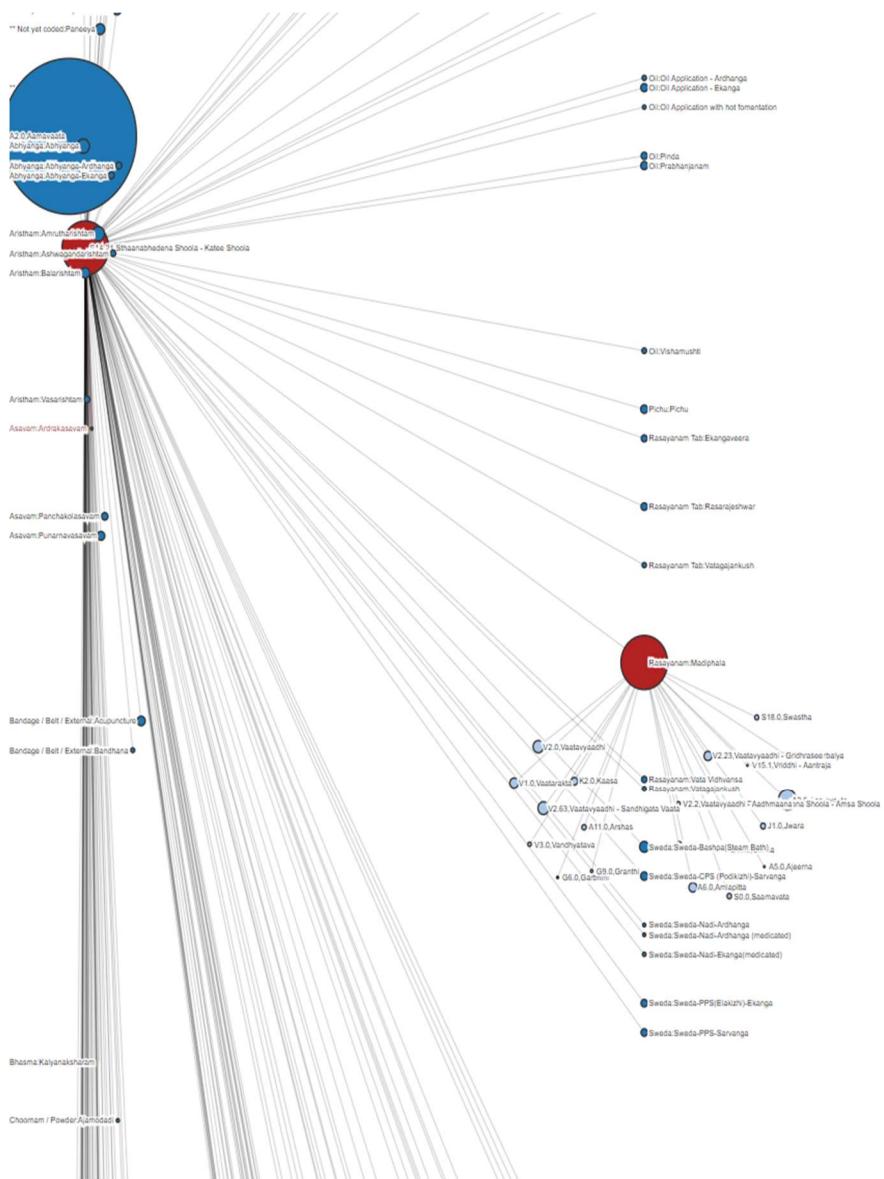


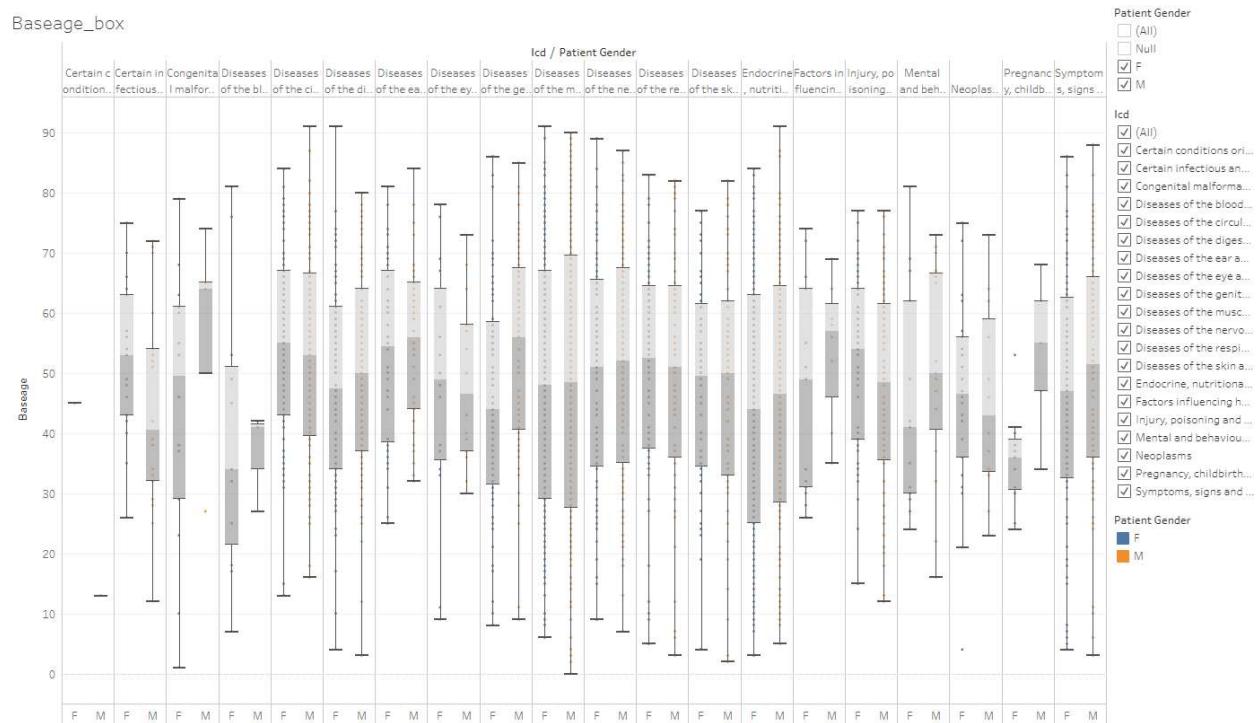
Figure 3-17: ICD classification by Gender

## ICDFreq

Icd	High	Grand Total	Patient Gender		
			Null	F	M
<b>Grand Total</b>		8,027	1	4,455	3,571
<b>Certain conditions originating in the perinatal period</b>	Total	2		1	1
	P91	2		1	1
<b>Certain infectious and parasitic diseases</b>	Total	90		42	48
	A01	9		6	3
	A06	1		1	
	A15	7		3	4
	A18	3			3
	A30	1		1	
	A41	1		1	
	A46	1		1	
	A81	2		2	
	A90	2		2	
	A92	5		4	1
	B01	1		1	
	B02	5			5
	B07	8		5	3
	B08	2		1	1
	B19	6			6
	B34	1			1
	B35	32		12	20
	B37	4		4	
	B74	1		1	
	B82	1			1

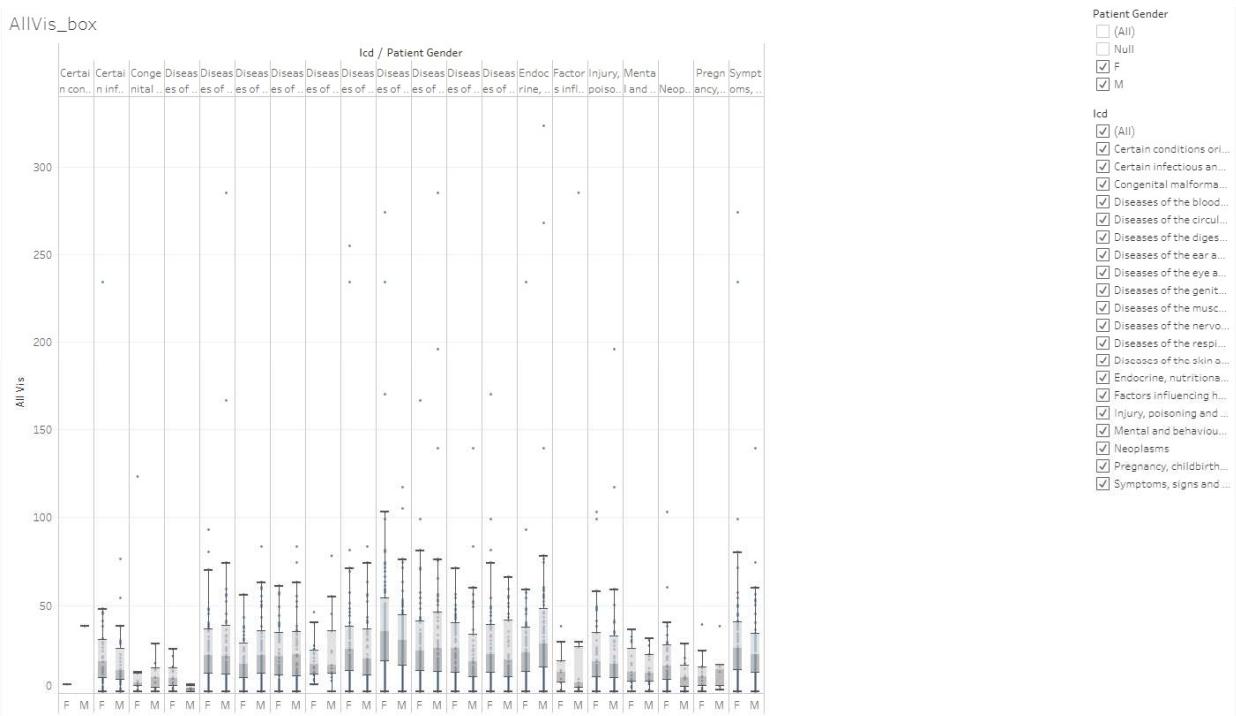
ICD classification and ICD classification high level categories, frequency counts by gender. Data version: 2011 to Oct 2017.  
Link to analysis: [Link](#)

Figure 3-18: Age distribution by ICD classification and Gender



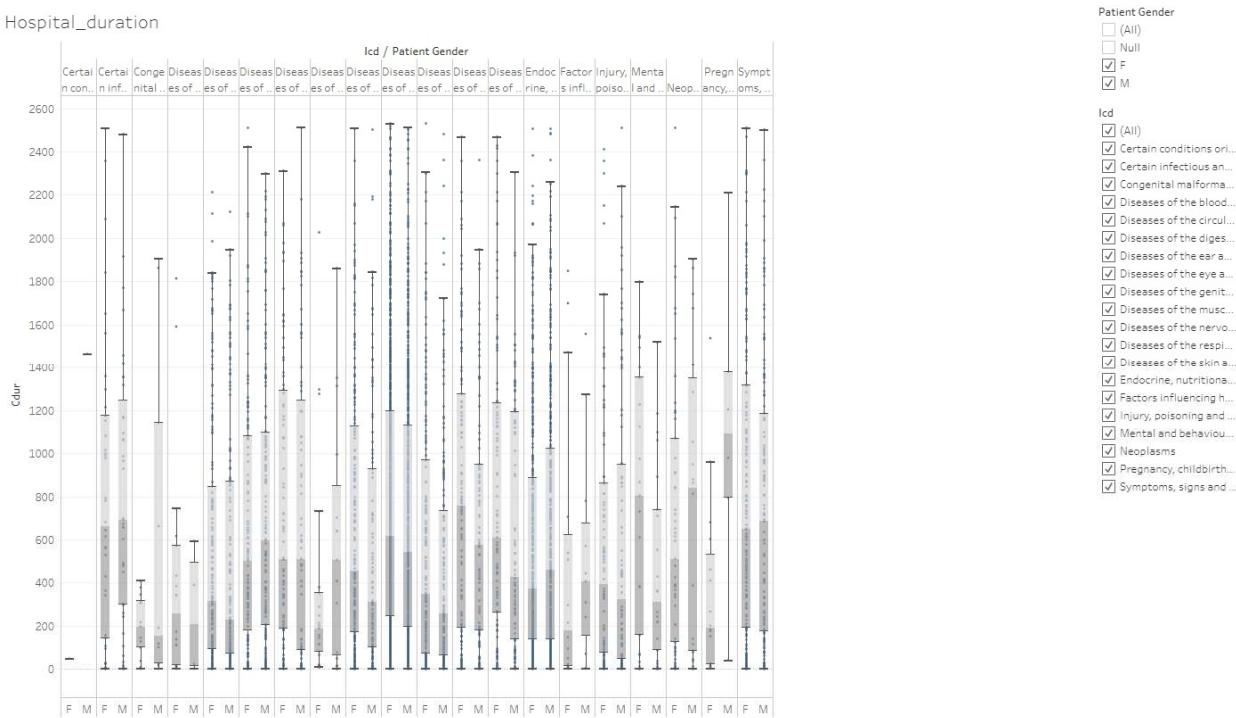
Boxplot representation of age by ICD classification and gender, Data version: 2011 to Oct 2017. Link to analysis: [Link](#)

Figure 3-19: Visit distribution by ICD classification and Gender



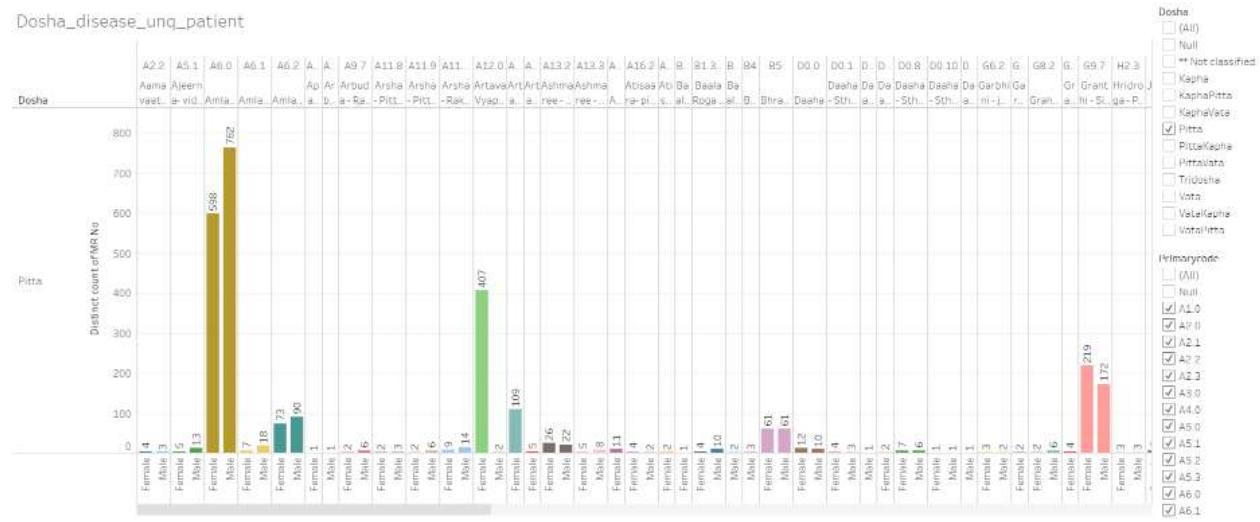
Boxplot representation of hospital number of visits by ICD classification and gender, Data version: 2011 to Oct 2017. Link to analysis: [Link](#)

Figure 3-20: Duration distribution by ICD classification and Gender



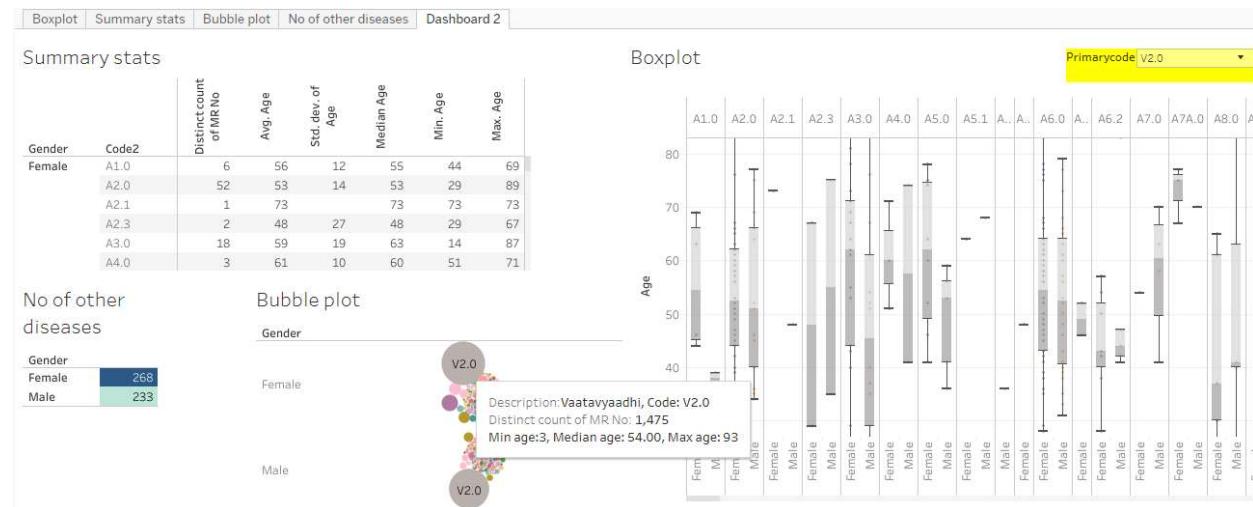
Boxplot representation of hospital visit duration by ICD classification and gender, Data version: 2011 to Oct 2017. Link to analysis: [Link](#)

Figure 3-21: Disease classification by Prakriti and Gender



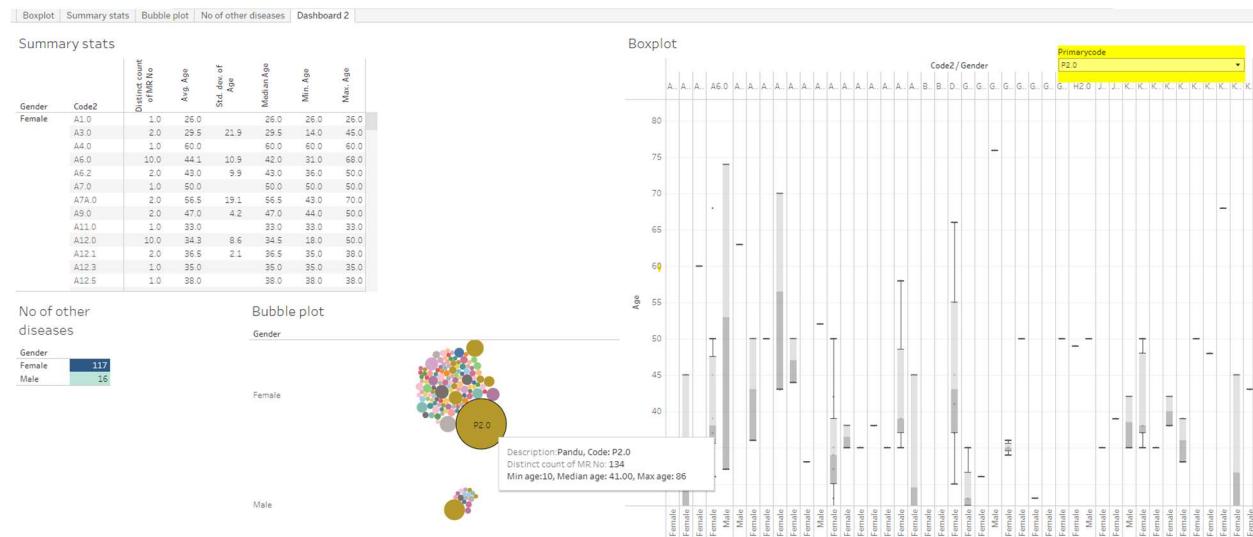
Dosha: Prakriti type, x-axis: male and female grouped by individual disease, y-axis: frequency counts of unique patients, Data version: 2011 to Oct 2016. Link to analysis: [Link](#) to get the above display subset Dosha for Pitta.

Figure 3-22: Co-morbidity analysis approach 1 example 1: Vaatavyadhi



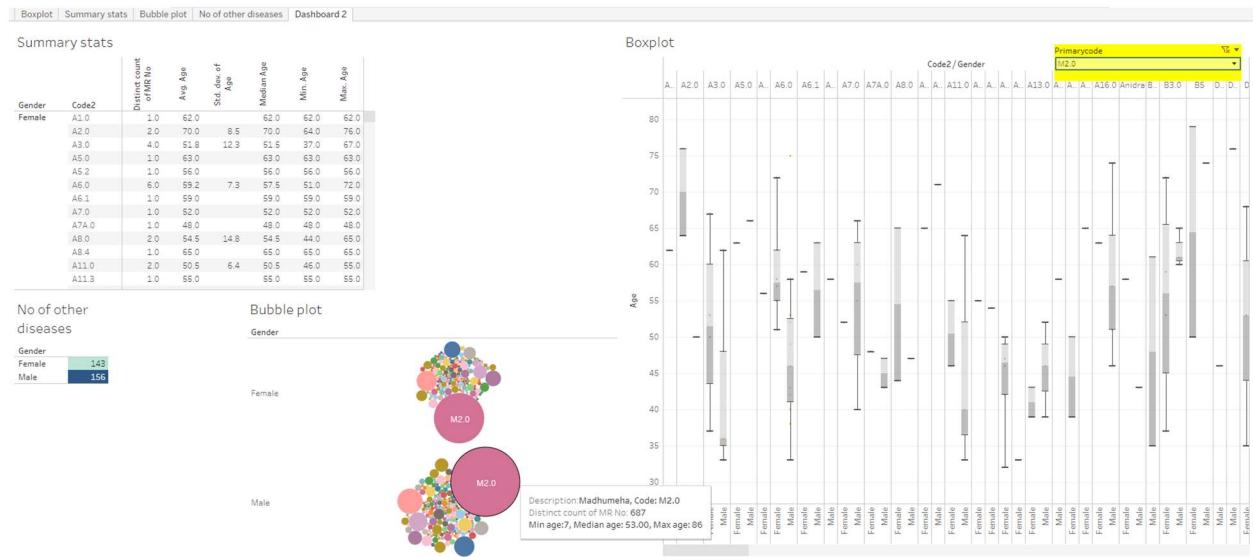
Summary stats section: descriptive statistics details by gender and other diseases reported, No of other disease: distinct number of other diseases reported by patients who had reported the primary disease, Bubble plot: frequency count of distinct patients by disease, Boxplot: age distribution by disease and gender, Data version: 2011 to Oct 2016. Link to analysis: [Link](#) to get the above display subset Primarycode = V2.0

Figure 3-23: Co-morbidity analysis approach 1 example 2: Pandu



Summary stats section: descriptive statistics details by gender and other diseases reported, No of other disease: distinct number of other diseases reported by patients who had reported the primary disease, Bubble plot: frequency count of distinct patients by disease, Boxplot: age distribution by disease and gender, Data version: 2011 to Oct 2016. Link to analysis: [Link](#) to get the above display subset Primarycode = P2.0

Figure 3-24: Co-morbidity analysis approach 1 example 3: Madhumeha



Summary stats section: descriptive statistics details by gender and other diseases reported, No of other disease: distinct number of other diseases reported by patients who had reported the primary disease, Bubble plot: frequency count of distinct patients by disease, Boxplot: age distribution by disease and gender, Data version: 2011 to Oct 2016. Link to analysis: [Link](#) to get the above display subset Primarycode = M2.0

Figure 3-25: Co-morbidity analysis approach 2



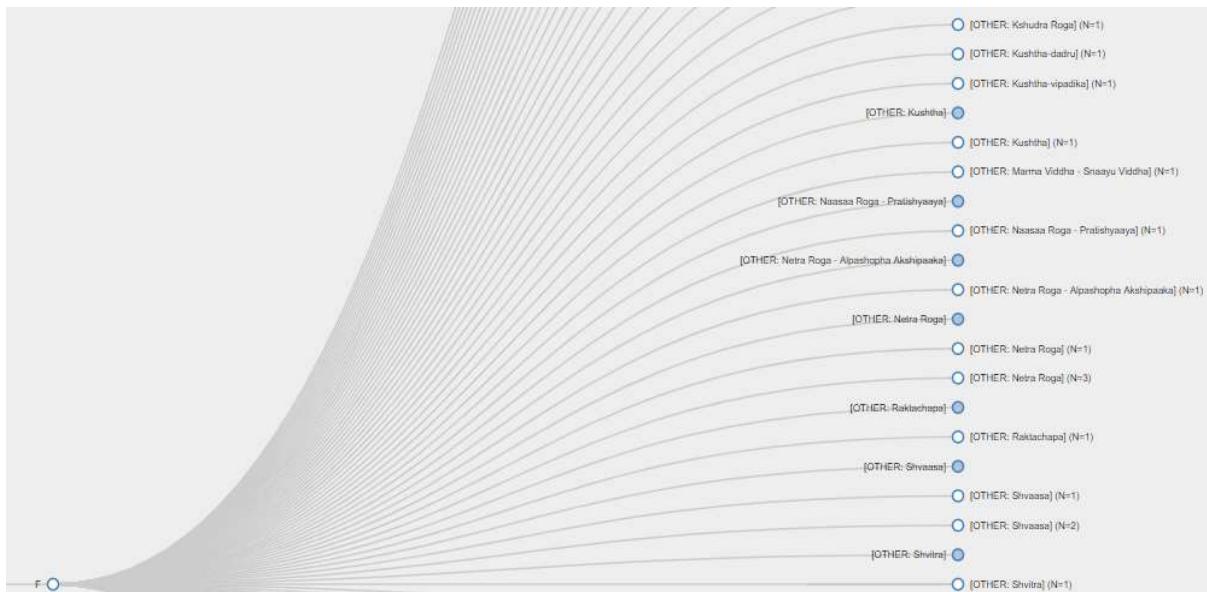
Upper section: bubble plots for the reference disease and other diseases reported, bubble size is based on number of distinct patients. Lower sections: unique number of other diseases reported for the reference disease. 1, 2, ..., 12: January to December month, Data version: 2011 to Oct 2016. Link to analysis: [Link](#) to get the above display subset Primarycode = V2.0

Figure 3-26: Co-morbidity analysis approach 3: collapsible tree view

Initial view of the tree



After clicking on F (Female), the collapsible tree opens up



An example of a disease experienced only by one gender



<https://coursephd.github.io/>

Figure 3-27: Patient Disease and Treatment administration by Study Day

Patient_Dis_Med_counts							Mr No
Mr No	Grpday	Studyday	Cat / Newold2		Medicine		Newold2
			Disease	1sttim..	Repeat	1st dose	
MR000335	1	1	1		4		1st dose
	2	24		1	1	2	1st time disease
	3	69		1	1		Repeat
	4	521	1		1		
	5	565		1	2	1	
	6	572	5		10		
	7	646	1		3		
	8	1093	1		1	2	

Patient_1st_Repeat				
Mr No	Grpday	Cat	Type med	Coded med
MR000335	1	Disease	** Not yet coded	** Not yet coded
		Medicine	Asavam	Lodhrasavam
			Kashayam	Varanadi
				Vasaguloochyadi
		Tablet / Gulika / Vati	Kanchanara Guggulu	
	2	Disease	** Not yet coded	** Not yet coded
		Medicine	Asavam	Lohasavam

The report displays individual patient data. Upper left part: Mr No: Patient ID, study day, Disease reported 1<sup>st</sup> time and repeated, medicine reported 1<sup>st</sup> time and repeated, Lower part of the report displays individual patient data for each day and distinguishes 1<sup>st</sup> dose, 1<sup>st</sup> disease and Repeat reporting for the same. Data version: 2011 to Oct 2017. Link to analysis: [Link](#) to get the above display subset Mr No = MR000335

Figure 3-28: Patient Disease by Study Day and Treatment administration by Study Day

Disease					Medicine				
Mr No	Grpday	Studyday	Newold2Dis		Mr No	Grpday	Studyday	Newold2	
			1st time	disease				1st	Repeat
MR000335	1	1		1	MR000335	1	1	4	
	2	24		1		2		1	2
	3	69		1		3	69	1	
	4	521		1		4	521	1	
	5	565		1		5	565	2	1
	6	572		1		6	572	10	
	7	646		1		7	646	3	
	8	1093		1		8	1093	1	2
	9	1388		1		9	1388	3	
	10	1418		1		10	1418	2	2
	11	1481		1		11	1481	4	
	12	1528		1		12	1528	3	
	13	1628		1		13	1628	2	1
	14	2025		2		14	2025	6	1
	15	2488		1		15	2488	1	
	16	2503		1		16	2503	2	
	17	2508		2		17	2508	2	1

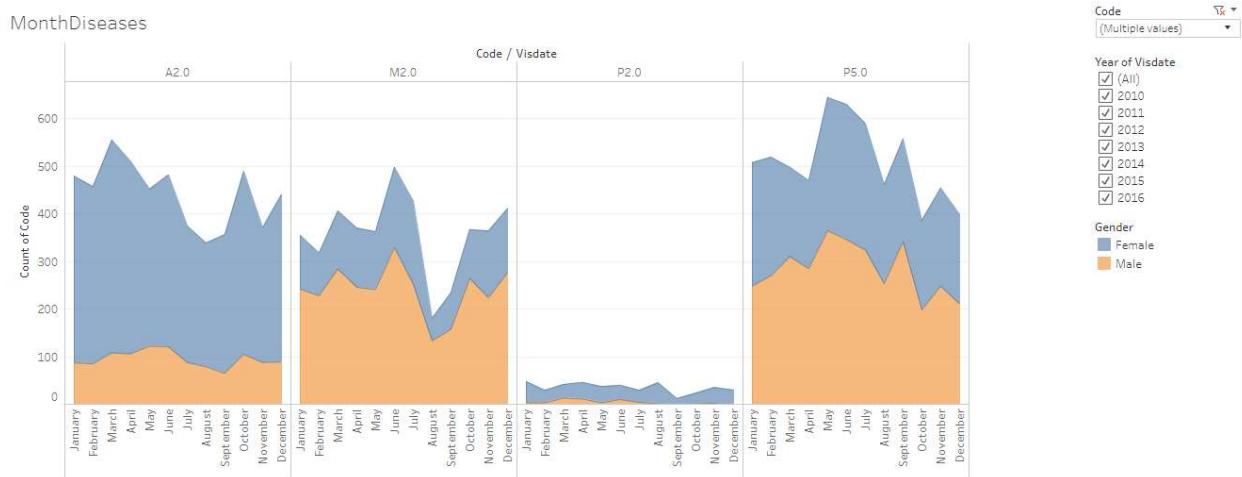
The report displays individual patient data. Panel on the left hand: number of diseases reported at a particular visit, panel on the right hand: number of prescribed treatments reported at a particular visit, Data version: 2011 to Oct 2017. Link to analysis: [Link](#) to get the above display subset Mr No = MR000335

Figure 3-29: Patient Cumulative Disease and Treatment administration by Visit

Patient_dis_med_Cumulative				%Patient_dis_med_Cumulative			
Mr No	Cumday2	Cat / Newold2		Mr No	Cumday2	Cat / Newold2	
		Disease	Medicine			Disease	Medicine
MR000335	Till visit 1	1	4	MR000335	Till visit 1	100.00%	100.00%
	Till visit 2	1	5		Till visit 2	50.00%	50.00%
	Till visit 3	1	6		Till visit 3	33.33%	66.67%
	Till visit 4	2	6		Till visit 4	50.00%	50.00%
	Till visit 5	2	8		Till visit 5	40.00%	60.00%
	Till visit 6	7	18		Till visit 6	70.00%	30.00%
	Till visit 7	8	21		Till visit 7	72.73%	27.27%
	Till visit 8	9	22		Till visit 8	75.00%	25.00%
	Till visit 9	10	25		Till visit 9	76.92%	23.08%
	Till visit 10	10	27		Till visit 10	71.43%	28.57%
	Till visit 11	11	31		Till visit 11	73.33%	26.67%
	Till visit 12	12	34		Till visit 12	75.00%	25.00%
	Till visit 13	13	36		Till visit 13	76.47%	23.53%
	Till visit 14	15	42		Till visit 14	78.95%	21.05%
	Till visit 15	16	43		Till visit 15	80.00%	20.00%
	Till visit 16	17	45		Till visit 16	80.95%	19.05%

The report displays individual patient data. Panel on the left hand: absolute values of diseases and prescribed treatments till particular visits, panel on the right hand side: % values of diseases and prescribed treatments till particular visits, Data version: 2011 to Oct 2017. Link to analysis: [Link](#)

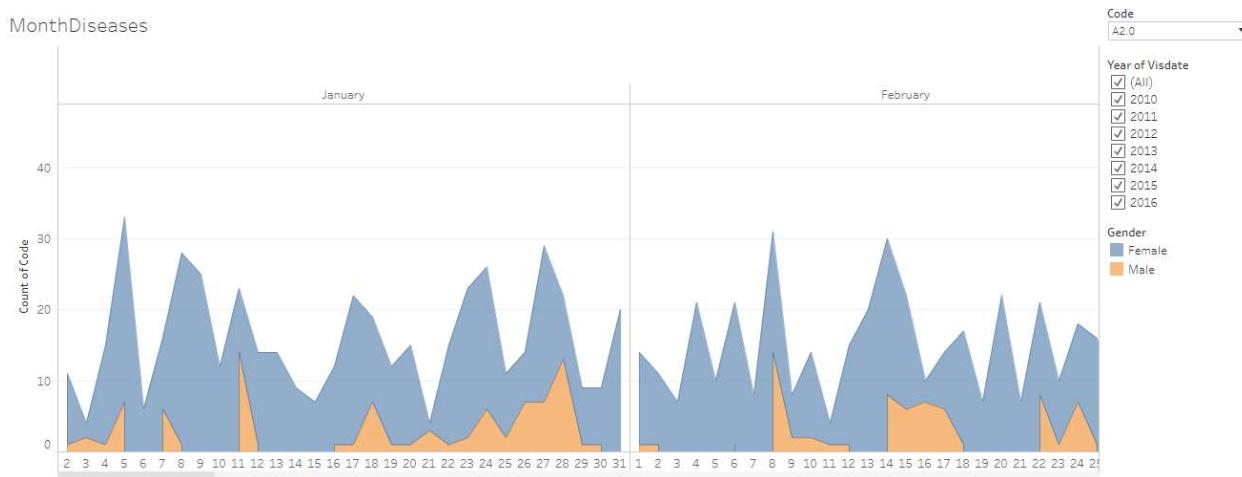
Figure 3-30: Area graph representation of diseases



Area graph representation of diseases: x-axis: Month, y-axis: frequency count of unique patients, disease code: the underlying data for each disease, peach colour: counts for male, blue colour: counts for female.

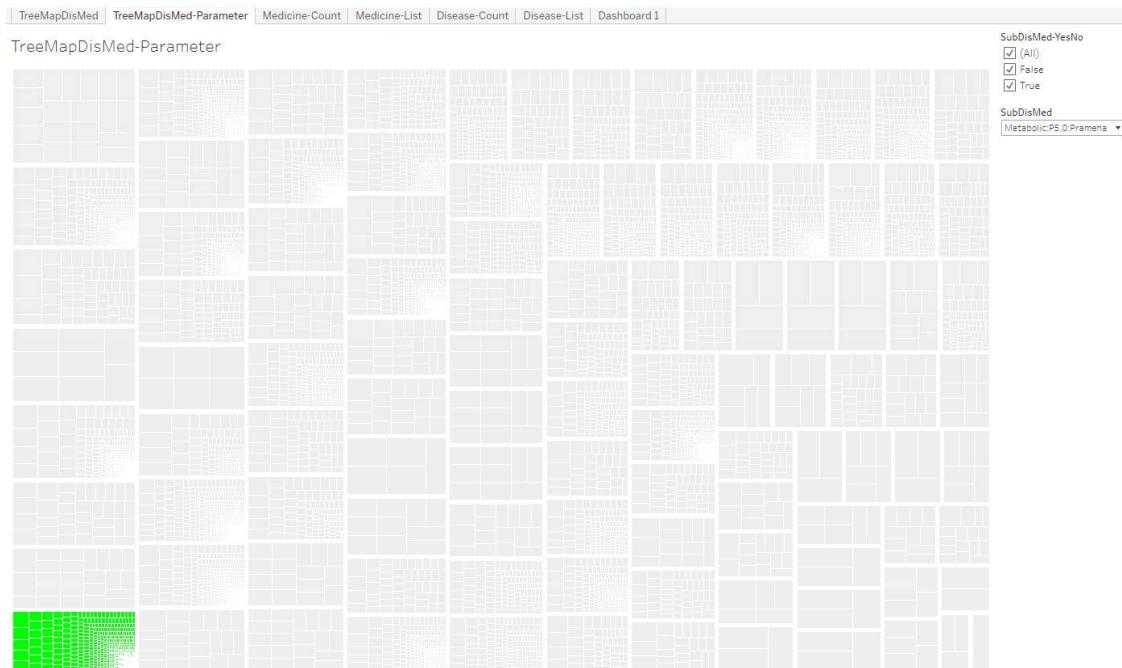


x-axis can be expanded to convert the monthly view to more granular unit (week, day).

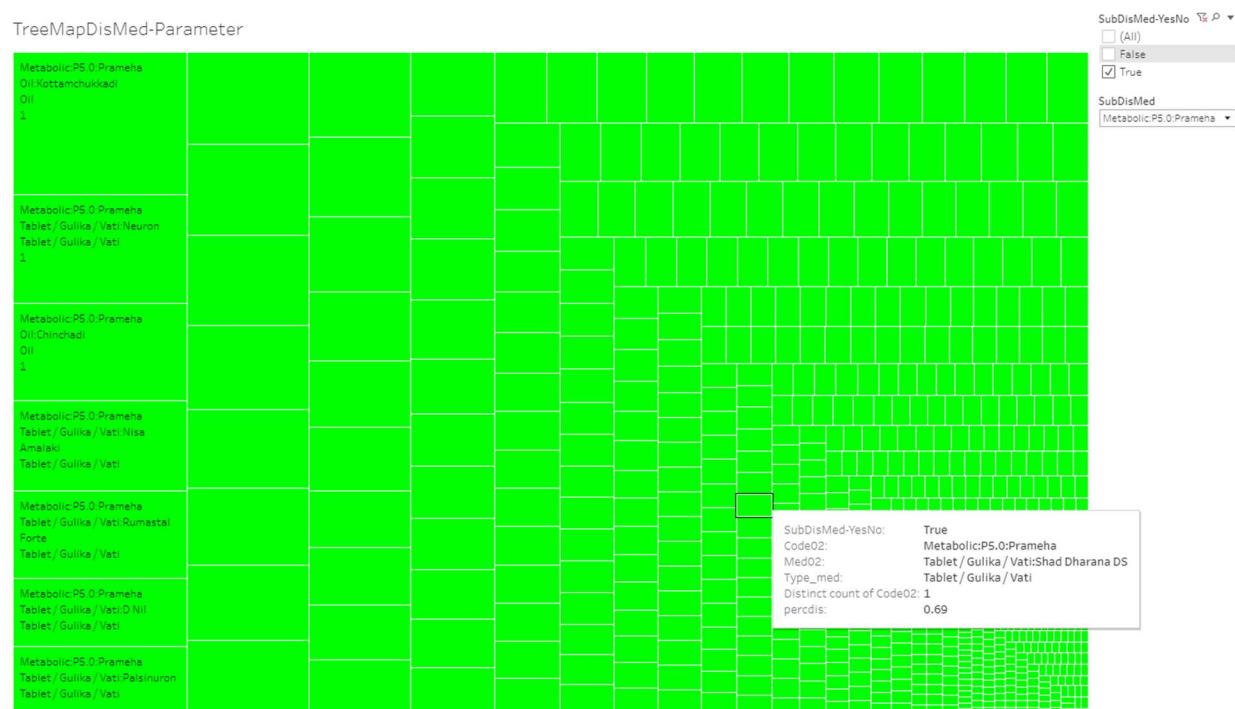


The original data displayed in the 1<sup>st</sup> part of the presentation is opened for daily view for a particular disease code A2.0, Data version: 2011 to 2016. Link to analysis: [Link](#) to get the above display subset the Code for (A2.0, M2.0, P2.0, P5.0)

Figure 3-31: Mosaic plot: Disease and treatment representation example 1: Prameha



Mosaic plot: Each box is one disease, the selected disease is marked in Green colour, smaller boxes inside each disease display one intervention each, Data version: 2011 to Oct 2017. Link to analysis: [Link](#) to get the above display subset SubDisMed = Metabolic: P5.0: Prameha



The snapshot above is the zoomed version of the Prameha block which highlights one of the many treatments prescribed for the same.

## Medicine Count

732

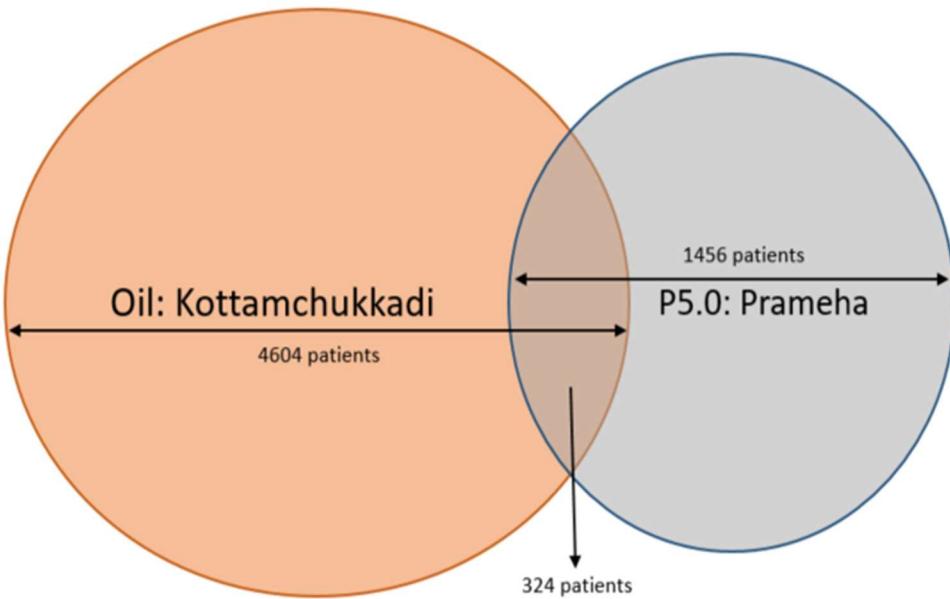
The Medicine count tab provides the information on total number of different medicines prescribed for Prameha. There are 732 distinct interventions.

Medicine-List							SubDisMed	
Med02	Code02	totpatmed	totpatdis	n	F	percdis	percmed	SubDisMed
Oil:Kottamchukkadi	Metabolic:P5.0:Prameha	4604	1456	324	22.25	7.04		Abc
Tablet /Gulika / Vati:Neuron	Metabolic:P5.0:Prameha	2065	1456	247	16.96	11.96		Abc
Oil:Chinchadi	Metabolic:P5.0:Prameha	2947	1456	221	15.18	7.5		Abc
Tablet /Gulika / Vati:Nisa Amalaki	Metabolic:P5.0:Prameha	798	1456	206	14.15	25.81		Abc
Tablet /Gulika / Vati:Rumastal Forte	Metabolic:P5.0:Prameha	1844	1456	199	13.67	10.79		Abc
Tablet /Gulika / Vati:D Nil	Metabolic:P5.0:Prameha	549	1456	155	10.65	28.23		Abc
Tablet /Gulika / Vati:Palsimuron	Metabolic:P5.0:Prameha	1511	1456	154	10.58	10.19		Abc
Tablet /Gulika / Vati:Diabecon DS	Metabolic:P5.0:Prameha	476	1456	145	9.96	30.46		Abc
Liniment:Myostal	Metabolic:P5.0:Prameha	1991	1456	145	9.96	7.28		Abc
Tablet /Gulika / Vati:Cervilon	Metabolic:P5.0:Prameha	1151	1456	143	9.82	12.42		Abc
Kashayam:RasnasaPthakam	Metabolic:P5.0:Prameha	1866	1456	133	9.13	7.13		Abc
Kashayam:Prasaranayadi	Metabolic:P5.0:Prameha	974	1456	124	8.52	12.73		Abc
Kashayam:Rasnærandadi	Metabolic:P5.0:Prameha	2265	1456	122	8.38	5.39		Abc
Tablet /Gulika / Vati:Nuro - XT	Metabolic:P5.0:Prameha	1026	1456	120	8.24	11.7		Abc
Kashayam:Rasnayograj Guggulu	Metabolic:P5.0:Prameha	967	1456	117	8.04	12.1		Abc
Kashayam:Asanadi	Metabolic:P5.0:Prameha	418	1456	113	7.76	27.03		Abc
Tablet /Gulika / Vati:Rumalaya Forte	Metabolic:P5.0:Prameha	1175	1456	106	7.28	9.02		Abc
Kashayam Tab:Rasnærandadi	Metabolic:P5.0:Prameha	1182	1456	100	6.87	8.46		Abc
Kashayam:Sahacharadi	Metabolic:P5.0:Prameha	1603	1456	92	6.32	5.74		Abc
Tablet /Gulika / Vati:Trayodasanga Guggulu	Metabolic:P5.0:Prameha	1830	1456	88	6.04	4.81		Abc
Tablet /Gulika / Vati:Lumbaton	Metabolic:P5.0:Prameha	1433	1456	84	5.77	5.86		Abc
Tablet /Gulika / Vati:Ksheerabala	Metabolic:P5.0:Prameha	594	1456	76	5.22	12.79		Abc
Kashayam:Ashtavargam	Metabolic:P5.0:Prameha	548	1456	76	5.22	13.87		Abc
Tablet /Gulika / Vati:Flexy	Metabolic:P5.0:Prameha	562	1456	73	5.01	12.99		Abc
Oil:Narayana	Metabolic:P5.0:Prameha	374	1456	73	5.01	19.52		Abc
Oil:Murivenna	Metabolic:P5.0:Prameha	2329	1456	69	4.74	2.96		Abc
Oil:Sahacharadi	Metabolic:P5.0:Prameha	1100	1456	67	4.6	6.09		Abc
Kashayam:Maharasnayograj Guggulu	Metabolic:P5.0:Prameha	660	1456	67	4.6	10.15		Abc
Tablet /Gulika / Vati:Chandraprabha	Metabolic:P5.0:Prameha	612	1456	66	4.53	10.78		Abc
Tablet /Gulika / Vati:Trayodasanga Guggulu DS	Metabolic:P5.0:Prameha	1249	1456	65	4.46	5.2		Abc
Tablet /Gulika / Vati:Mehamil	Metabolic:P5.0:Prameha	214	1456	59	4.05	27.57		Abc
Tablet /Gulika / Vati:Glukostat	Metabolic:P5.0:Prameha	175	1456	57	3.91	32.57		Abc
Oil:Pinda	Metabolic:P5.0:Prameha	937	1456	57	3.91	6.08		Abc
Rasayanam:Vasant Kusumakar Plain	Metabolic:P5.0:Prameha	535	1456	55	3.78	10.28		Abc
Tablet /Gulika / Vati:Sallaki M R	Metabolic:P5.0:Prameha	2972	1456	54	3.71	1.82		Abc
Kashayam:Kathaka Kathiradi	Metabolic:P5.0:Prameha	167	1456	53	3.64	31.74		Abc
Tablet /Gulika / Vati:Dia P	Metabolic:P5.0:Prameha	134	1456	51	3.5	38.06		Abc

A detailed list of medicines with the total patients prescribed with them as well as total patients suffering from Prameha are displayed in the snapshot above.

Figure 3-32: Disease and treatment example 2: P5.0: Prameha and Oil: Kottamchukkadi

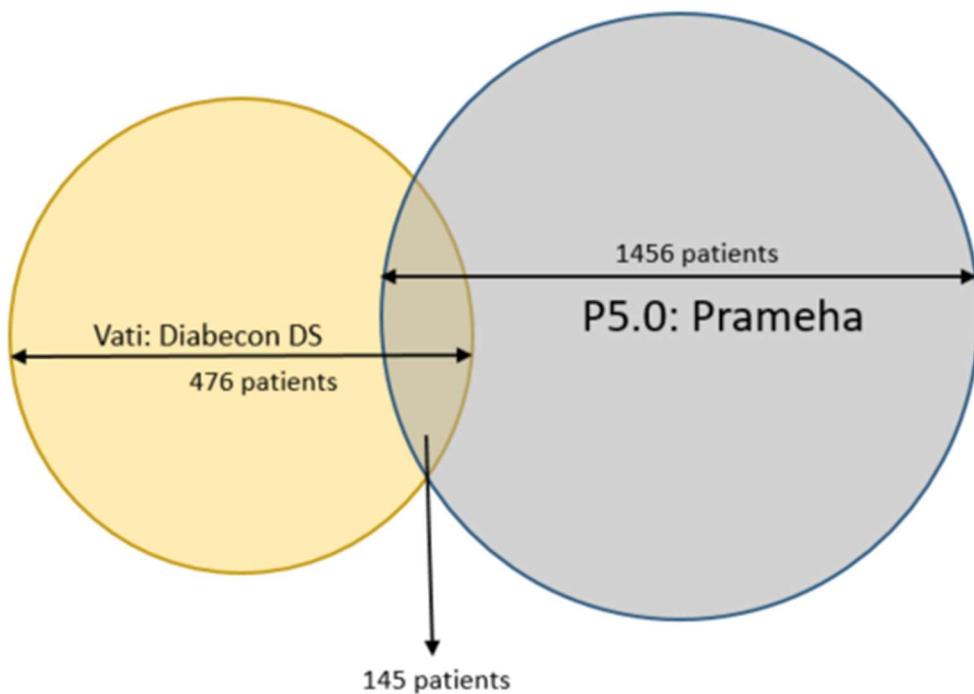
Medicine-List							
Med02	Code02	totpatmed	totpatdis	n	F	percdis	percmed
Oil:Kottamchukkadi	Metabolic:P5.0:Prameha	4604	1456	324	22.25	7.04	



Totpatmed shows that there are 4,604 patients who have been prescribed with Oil: Kottamchukkadi; Totpatdis shows that there are 1,456 patients who are diagnosed with Prameha; this shows that the treatment is not Prameha specific. The n count of 324 shows the patients who had Prameha and were prescribed Oil: Kottamchukkadi. Percdis shows that 22.25% i.e. 324 / 1456 of patients having Prameha are prescribed this particular treatment. Percmed shows that only 7% of the time this medicine has been prescribed for Prameha patients.

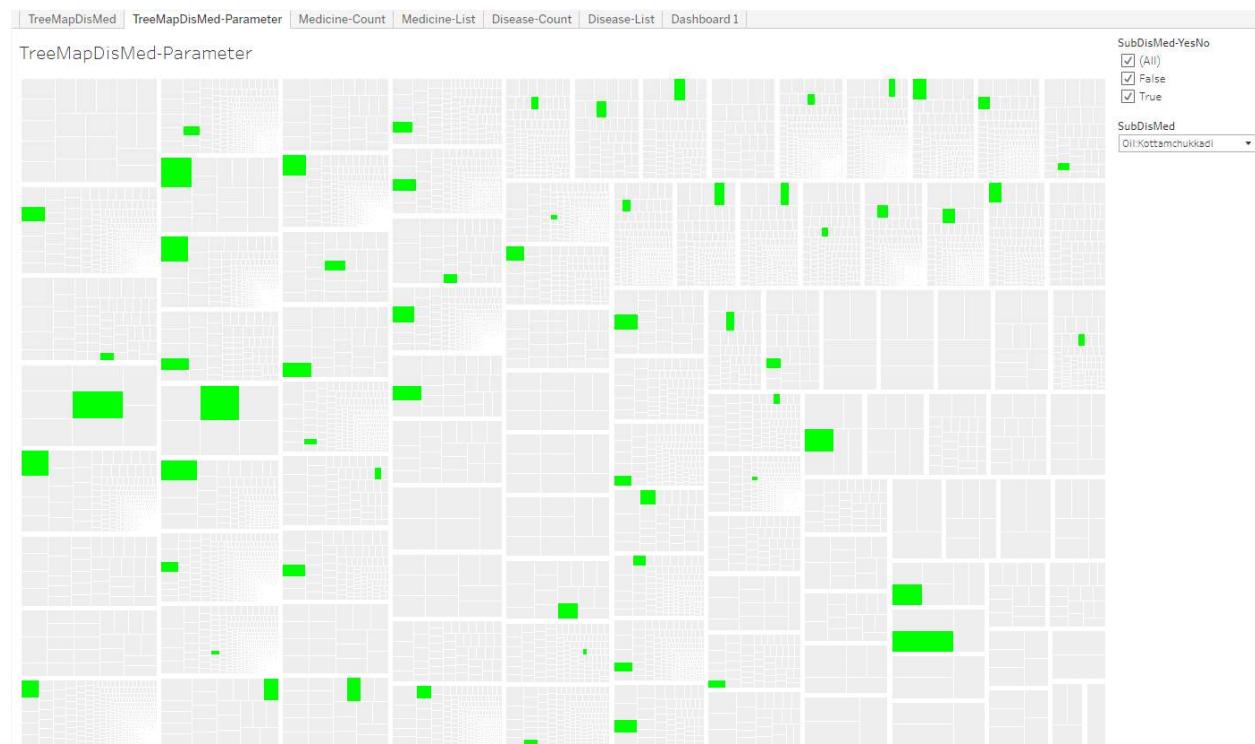
Figure 3-33: Disease and treatment example 3: P5.0: Prameha and Vati: Diabecon DS

Tablet / Gulika / Vati:Diabecon DS	Metabolic:P5.0:Prameha	476	1456	145	9.96	30.46
------------------------------------	------------------------	-----	------	-----	------	-------



Totpatmed shows that there are 476 patients who have been prescribed with Vati: Diabecon DS; Totpatdis shows that there are 1,456 patients who are diagnosed with Prameha; this shows that the treatment could be more prescribed to Prameha patients. The n count of 145 shows the patients who had Prameha and were prescribed Vati: Diabecon DS. Percdis shows that 9.96% i.e. 145 / 1,456 of patients having Prameha are prescribed this particular treatment. Percmed shows that only 30.46% of the time this medicine has been prescribed for Prameha patients.

Figure 3-34: Mosaic plot Disease and treatment representation example 4: Treatment: Oil: Kottamchukkadi



Mosaic plot: Each box is one disease, the selected disease is marked in Green colour, smaller boxes inside each disease display one intervention each, Data version: 2011 to Oct 2017. Link to analysis: [Link](#) to get the above display subset SubDisMed = Oil: Kottamchukkadi



The Disease count tab provides the information on total number of different diseases for which the intervention was prescribed. There are 62 distinct diseases.

Medicine-List							SubDisMed		
Med02	F	Code02	totpatmed	totpatdis	n	F	percdis	percmed	
OilKottamchukkadi		RMSD.V2.63:Vaataavyaad..	4604	4160	1506	36.2	32.71		Abc
		RMSD.V2.23:Vaataavyaad..	4604	2476	1274	51.45	27.67		Abc
		RMSD.S13.11:Sthaanabh..	4604	942	490	52.02	10.64		Abc
		RMSD.V2.0:Vaataavyaadhi..	4604	2822	396	14.03	8.6		Abc
		RMSD.S14.21:Sthaanabh..	4604	1227	326	26.57	7.08		Abc
		Metabolic:P5.0:Prameha..	4604	1456	324	22.25	7.04		Abc
		RMSD.A2.0:Amavaata..	4604	790	205	25.95	4.45		Abc
		RMSD.A3.0:Abhigataja S..	4604	703	78	11.1	1.69		Abc
		RMSD.S13.19:Sthaanabh..	4604	120	65	54.17	1.41		Abc
		RMSD.V1.0:Vaataarkta..	4604	506	62	12.25	1.35		Abc
		RMSD.V2.36:Vaataavyaad..	4604	361	51	14.13	1.11		Abc
		RMSD.S14.0:Sthaanabhe..	4604	508	50	9.84	1.09		Abc
		RMSD.V2.9:Vaataavyadhi..	4604	203	44	21.67	0.96		Abc
		Metabolic:M1.0.1:Medoro..	4604	590	42	7.12	0.91		Abc
		Metabolic:M2.0:Madhum..	4604	1395	40	2.87	0.87		Abc
		RMSD.S14.7:Sthaanabhe..	4604	227	35	15.42	0.76		Abc
		RMSD.S14.17:Sthaanabh..	4604	189	26	13.76	0.56		Abc
		RMSD.V2.75:Vaataavyaad..	4604	162	20	12.35	0.43		Abc
		RMSD.V2.30:Vaataavyaad..	4604	78	18	23.08	0.39		Abc
		Metabolic:S16.0:Sthulya..	4604	911	18	1.98	0.39		Abc
		RMSD.S15.42:Sthaanabh..	4604	73	15	20.55	0.33		Abc
		RMSD.S15.36:Sthaanabh..	4604	75	14	18.67	0.3		Abc
		RMSD.S15.31:Sthaanabh..	4604	86	13	15.12	0.28		Abc
		RMSD.V2.77:Vaataavyaad..	4604	55	11	20	0.24		Abc
		Metabolic:M1.0.0:Medoro..	4604	139	11	7.91	0.24		Abc
		Metabolic:P5.4:Prameha ..	4604	140	9	6.43	0.2		Abc
		RMSD.S14.3:Sthaanabhe..	4604	79	8	10.13	0.17		Abc
		RMSD.V2.68:Vaataavyaad..	4604	49	7	14.29	0.15		Abc
		RMSD.S14.24:Sthaanabh..	4604	29	7	24.14	0.15		Abc
		RMSD.S14.18:Sthaanabh..	4604	34	7	20.59	0.15		Abc
		RMSD.V2.47:Vaataavyaad..	4604	24	6	25	0.13		Abc
		RMSD.S10.4:Stambha -Gr..	4604	31	6	19.35	0.13		Abc
		RMSD.A2.3:Amavaasta ..	4604	30	5	16.67	0.11		Abc
		Metabolic:M1.0.2:Medoro..	4604	188	5	2.66	0.11		Abc
		RMSD.V2.42:Vaataavyaad..	4604	52	4	7.69	0.09		Abc
		RMSD.S14.4:Sthaanabhe..	4604	68	4	5.88	0.09		Abc
		RMSD.S14.13:Sthaanabh..	4604	45	3	6.67	0.07		Abc

A detailed list of diseases with the total patients prescribed with the treatment as well as total patients suffering from different diseases are displayed.

Figure 3-35: Cross tabulation of prescribed treatments and disease group by gender Example 1

### MedType\_DisType

Medicine Name	Distype / Patient Gender						Grand Total	
	Metabolic			RMSD				
	F	M	Total	F	M	Total		
Balarishtam				29	10	39	39	
Balarishtam 450 ml	3	2	5	143	80	223	225	
Balarishtam 450 ml [sna]				16	9	25	25	
Balarishtam 450 ml KPL				13	9	22	22	
Balarishtam 450 ML SNA				9	5	14	14	
Balarishtam 450 ml Vaidya	3	2	5	158	122	280	282	
Balarishtam [ KPL ]				2	1	3	3	
Balarishtam [AVP]	11	8	19	353	256	609	616	
Balarishtam Rajah				1	1	2	2	
Balarishtam(Arya)	1	1	2	24	12	36	37	
Grand Total	17	13	30	677	465	1,142	1,154	

Cross tabulation of medicine, disease type and patient gender, Medicine name: source data collected on Case Report Form, Distype: Metabolic and RMSD groups derived in the analysis dataset. Patient Gender: source data. Only 30 patients having metabolic diseases were prescribed the medicine whereas 1,142 patients with RMSD were prescribed. This reflects the ayurvedic principle of who should be prescribed any aristham. Data version: 2011 to Oct 2017. Link to analysis: [Link](#) to get the above display subset Medicine name = Balarishtam

Figure 3-36: Cross tabulation of prescribed treatments and disease group by gender Example 2

### MedType\_DisType

Medicine Name	Distype / Patient Gender						Grand Total
	Metabolic			RMSD		Total	
	F	M	Total	F	M	Total	
Abhrak bhasma (Dooth) 2gm				2	1	3	3
Abhraka Bhasma 1g (Sahasraputi) Dhoot					1	1	1
Abhraka Bhasma 5g [101] (AVP)	3	2	5	1	2	3	8
Abhraka Bhasma AVP	3	2	5	2	1	3	8
Abhraka Bhasma Dooth-100Butas	1	2	3	2	6	8	10
Abhrakam bhasma				2		2	2
Apamarga kshara Bhasma					1	1	1
Godhanti Bhasma 50gm [SDM]	1		1	16	4	20	20
Hajrulayahda Bhasma (Baidyanath)	1	5	6	3	8	11	17
Heeraka Bhasma 100mg [Dooth]				2		2	2
Kanmada Bhasma(10gm) (AVP)	4	4	8	2	2	4	11
Kashis Bhasma (Baidyanath)				1		1	1
Pravala bhasma	1		1	2	2	4	5
Pravala Bhasma Tab Dooth	1		1		1	1	2
Pravala Bhasmam (AVP)	1		1	1		1	2
Pravala Bhasmam Cap - 100				10	2	12	12
Pravala Bhasmam Cap (Kottakkal)	2		2	98	35	133	133
Rajatabhasma Cap (Kottakkal)	1		1				1
Roupya (Rajatha) Bhasma					1	1	1
Roupya (Rajatha) Bhasma 5 gm (Dooth)				2	1	3	3
Sameerapannaga Bhasma (Dooth)				2		2	2
sanka bhasmam				2		2	2
sankha bhasma				1	1	2	2
Sankha Bhasma(AVP)	1			1		1	1
Shilajith Bhasma	1	2	3				3
Shuddha Shilajit Bhasma		1	1				1
Tankan Bhasma 10gm		1	1	1	1	2	3
Tankan Bhasma (Baidyanath)		2	2	3	2	5	7
Tapyadi Loha Bhasma (Dooth)	1		1	3	1	4	4
Trivanga Bhasma (Dooth)	1	1	2				2
Vanga Bhasma 5g (Dhoot)	1	2	3				3
Yashada Bhasma 5gm (Dooth)	3	6	9	10	12	22	28
Grand Total	24	30	54	161	82	243	287

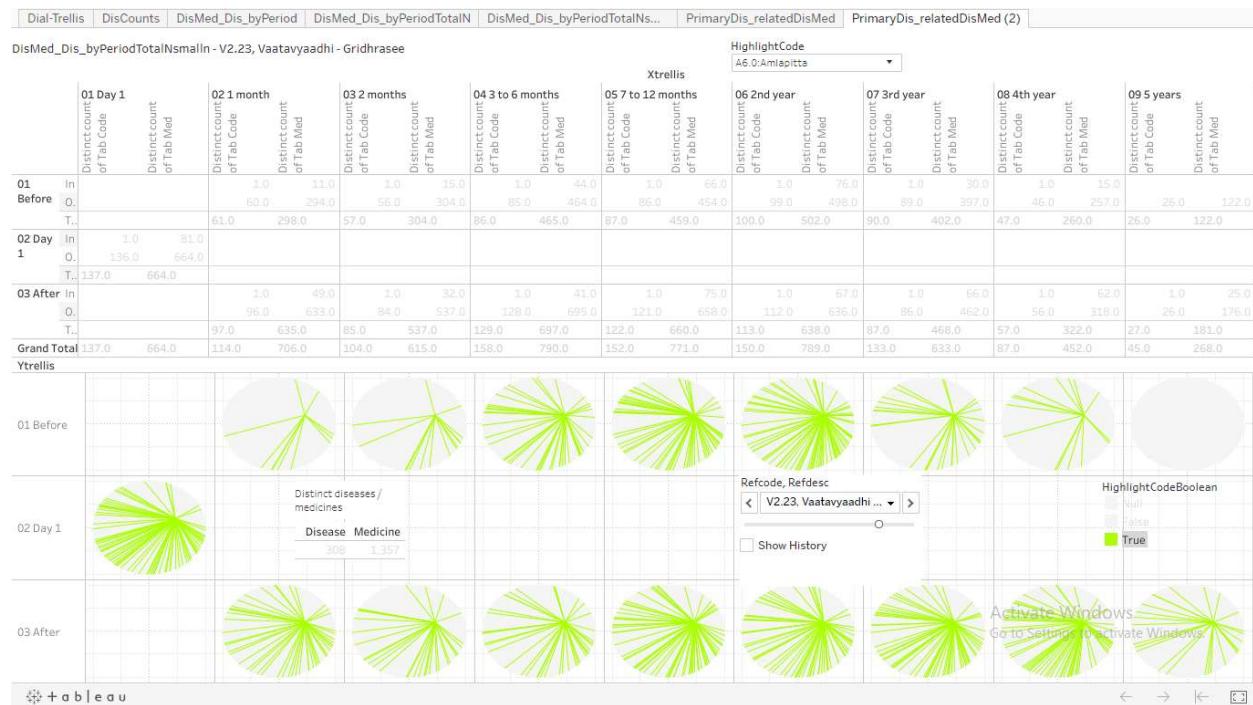
Cross tabulation of medicine, disease type and patient gender, Medicine name: source data collected on Case Report Form, Distype: Metabolic and RMSD groups derived in the analysis dataset. Patient Gender: source data. Only 287 (1.5%) patients have been prescribed bhasma. This reflects the ayurvedic principle of using bhasma based treatment wisely. Data version: 2011 to Oct 2016. Link to analysis: [Link](#) to get the above display subset Medicine name = Bhasma

Table 3-2: Summary statistics and t-test for bhasma usage

Overall (N=514)	Welch Two Sample t-test
<b>Pre_Bhasma</b>	
Mean (SD) 14.8 (14.0)	
Median [Min, Max] 9.00 [2.00, 111]	
<b>Post_Bhasma</b>	
Mean (SD) 10.6 (14.5)	
Median [Min, Max] 6.00 [0, 89.0]	

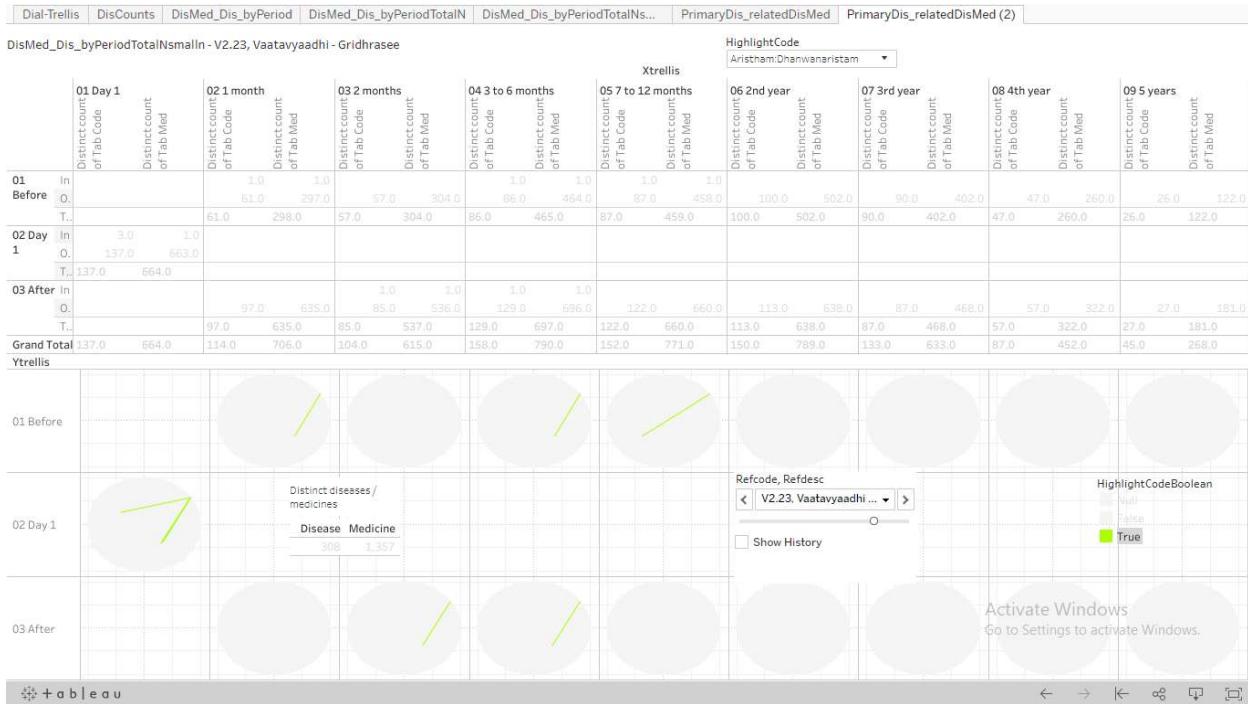
Patients who were prescribed at least one “bhasma” treatments are summarized, Pre\_Bhasma: Duration of treatment before the 1<sup>st</sup> bhasma treatment, Post\_Bhasma: Duration of treatment after the 1<sup>st</sup> bhasma treatment, Data version: 2011 to Oct 2017  
[https://github.com/coursephd/PostgreSQL/blob/master/110\\_rasa\\_aushadhi\\_analysis.R](https://github.com/coursephd/PostgreSQL/blob/master/110_rasa_aushadhi_analysis.R)

Figure 3-37: Circular view: Co-occurrences of disease – disease Example 1



Example 1: Disease: A6.0: Amavaata and Reference Disease: V2.23: Vaatavyadhi – Gridhrasee, Upper section: Pre and post time windows, count of distinct diseases and count of distinct medicines prescribed at the given time point. Lower section: 1<sup>st</sup> row represents the co-occurrence of disease – disease and / or disease – treatment before day 1 of the reference disease. Middle row: On day 1 count of distinct diseases and count of distinct medicines prescribed. Last row represents the same co-occurrence data after day 1 of the reference disease. Green bars inside a circle show co-occurrence of chosen disease – disease and / or disease – treatment combination, Data version: 2011 to Oct 2017. Link to analysis: [Link](#) to get the above display subset Refcode, Refdesc = V2.23 Vaatavyadhi – Gridhrasee and HighlightCode = A6.0: Amlapitta

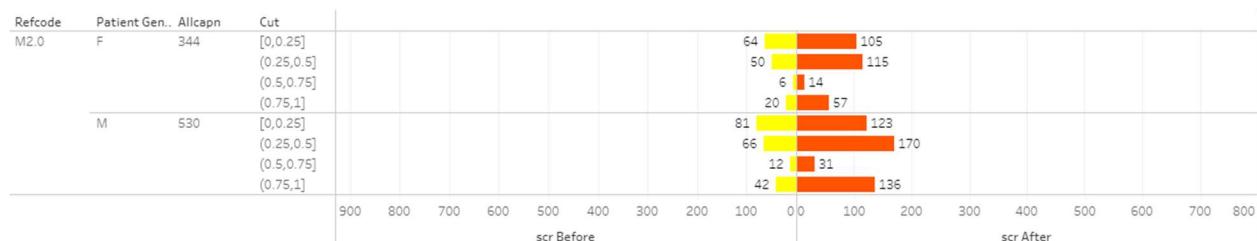
Figure 3-38: Circular view: Co-occurrences of disease – treatment Example 2



Example 2: Treatment: Arishtam:Dhanwanaristham and Reference Disease: V2.23: Vaatavyadhi – Gridhrasee, Upper section: Pre and post time windows, count of distinct diseases and count of distinct medicines prescribed at the given time point. Lower section: 1<sup>st</sup> row represents the co-occurrence of disease – disease and / or disease – treatment before day 1 of the reference disease. Middle row: On day 1 count of distinct diseases and count of distinct medicines prescribed. Last row represents the same co-occurrence data after day 1 of the reference disease. Green bars inside a circle show co-octene of chosen disease – disease and / or disease – treatment combination, Data version: 2011 to Oct 2017. Link to analysis: [Link](#) to get the above display subset Refcode, Refdesc = V2.23: Vaatavyadhi – Gridhrasee and HighlightCode = Arishtam:Dhanwanaristham

Figure 3-39: Pre and Post distance analysis for disease: M2.0: Madhumeha

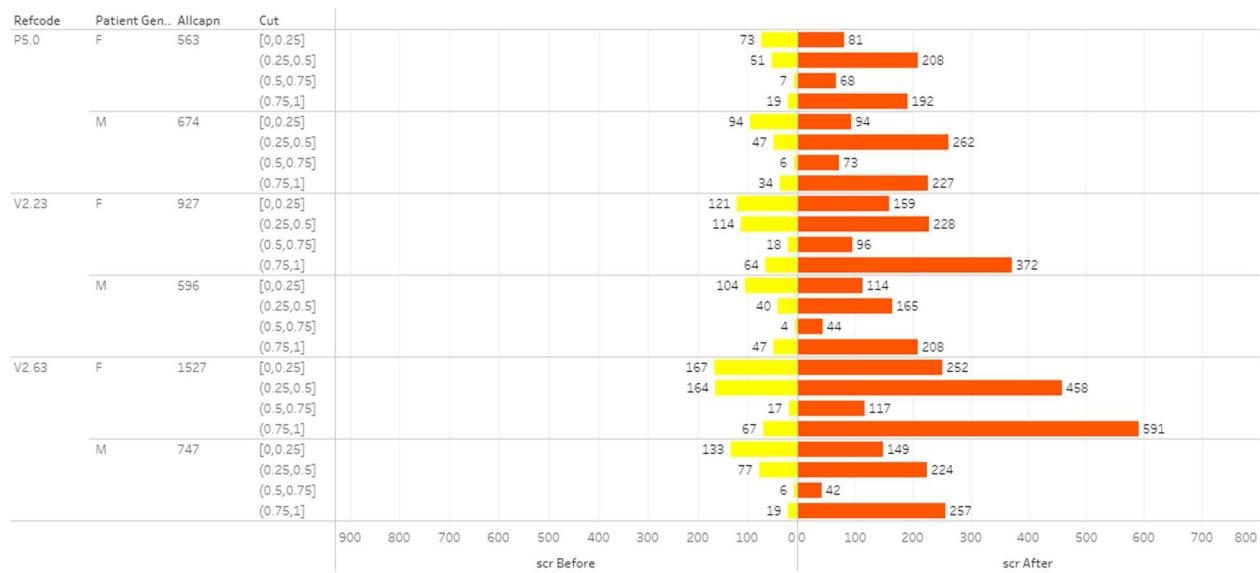
DiseaseMaxDist



Butterfly plot display of pre and post distance analysis for disease trajectories. Refcode: reference disease, Patient gender, Allcapn: Total number of patients in each of the categories, Cut: Jaccard distance, scr Before: patients falling in a particular category before day 1 of reference disease, scr After: patients falling in a particular category after day 1 of reference disease, Data version: 2011 to Oct 2017. Link to analysis: [Link](#) to get the above display subset Refcode = M2.0

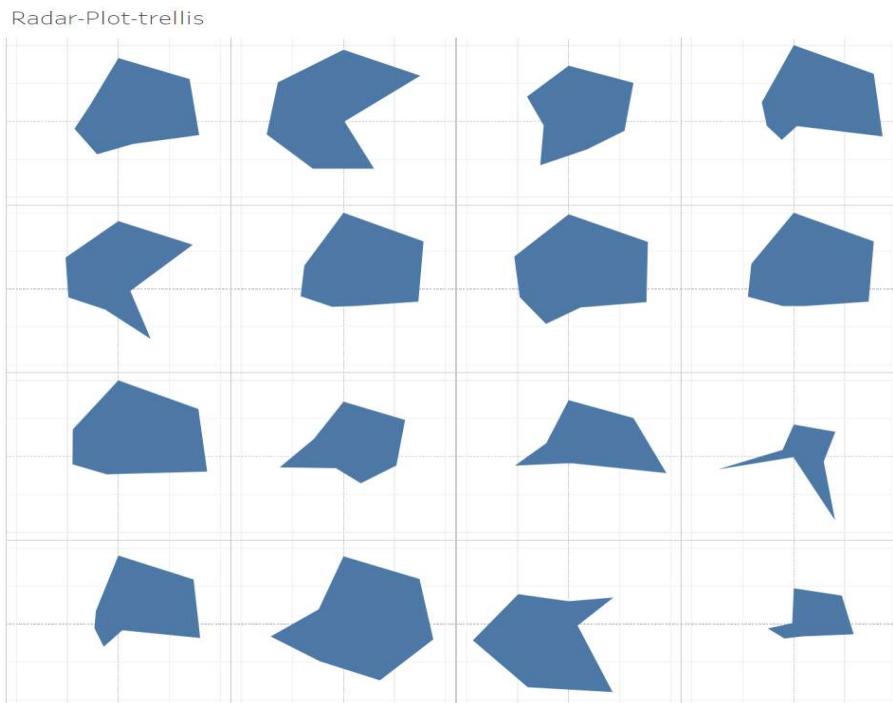
Figure 3-40: Pre and Post distance analysis for medicines given for diseases: P5.0, V2.23, V2.63

### DiseaseMaxDist



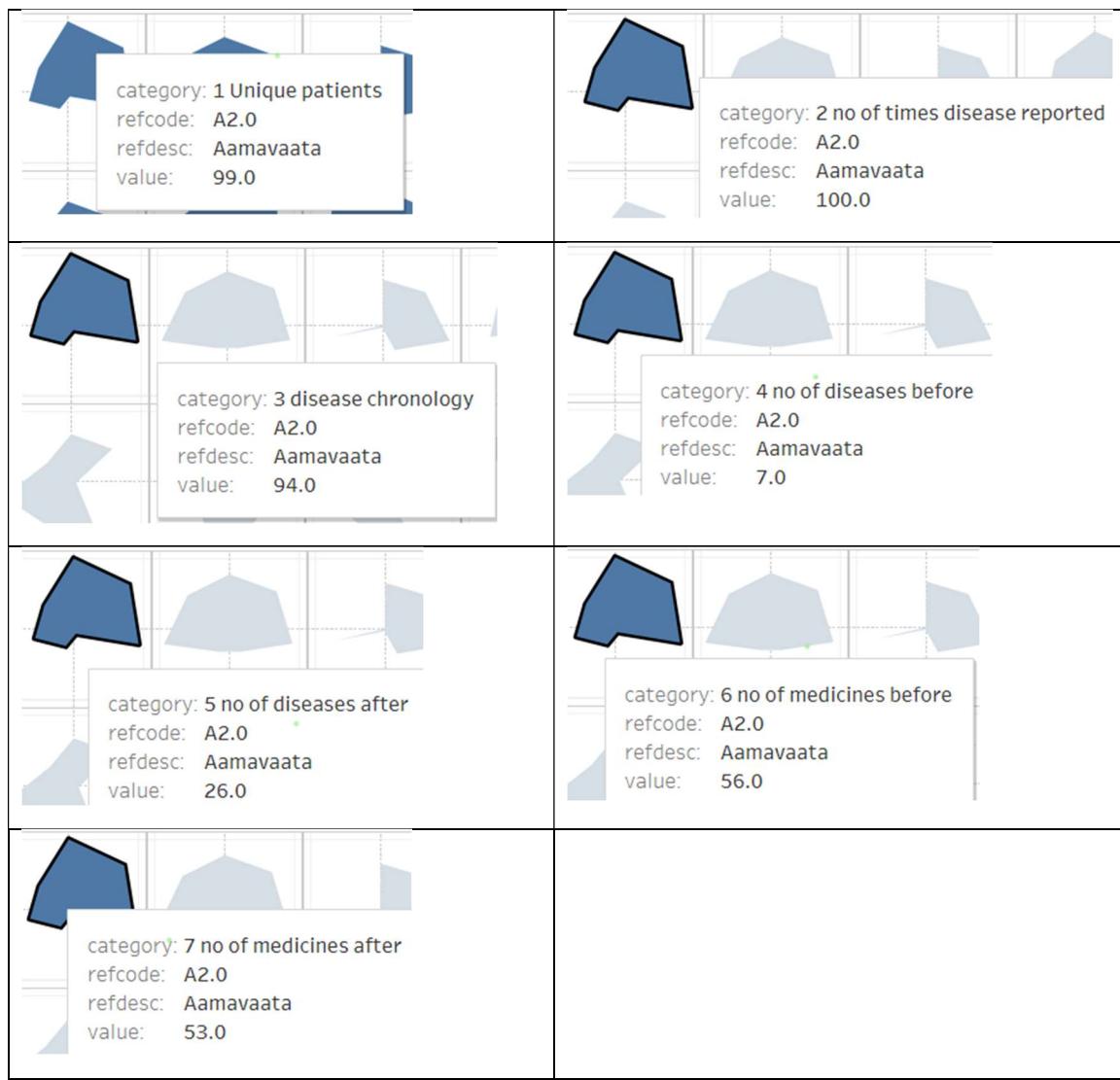
Butterfly plot display of pre and post distance analysis for medicine trajectories. Refcode: reference disease, Allcapn: Total number of patients in each of the categories, Cut: Jaccard distance, scr Before: patients falling in a particular category before day 1 of reference disease, scr After: patients falling in a particular category after day 1 of reference disease. This distance calculation was done on the basis of trajectory of prescribed treatments, Data version: 2011 to Oct 2017. Link to analysis: [Link](#) to get the above display subset for Refcode in (P5.0, V2.23, V2.63)

Figure 3-41: Radar plot



Radar plot showing multiple diseases displayed side by side.

Example 1: Multidimensional view a single disease: A6.0: Aamavaata



7 parameters displayed on one plot: (1) Unique patients, (2) Number of times disease reported, (3) Disease chronology, (4) Number of diseases reported before the reference disease, (5) Number of disease reported after the reference disease, (6) Number of medicines prescribed before the reference disease, (7) Number of medicines prescribed after the reference disease, Data version: 2011 to Oct 2017. Link to analysis: [Link](#)

Example 2: Multiple disease comparison

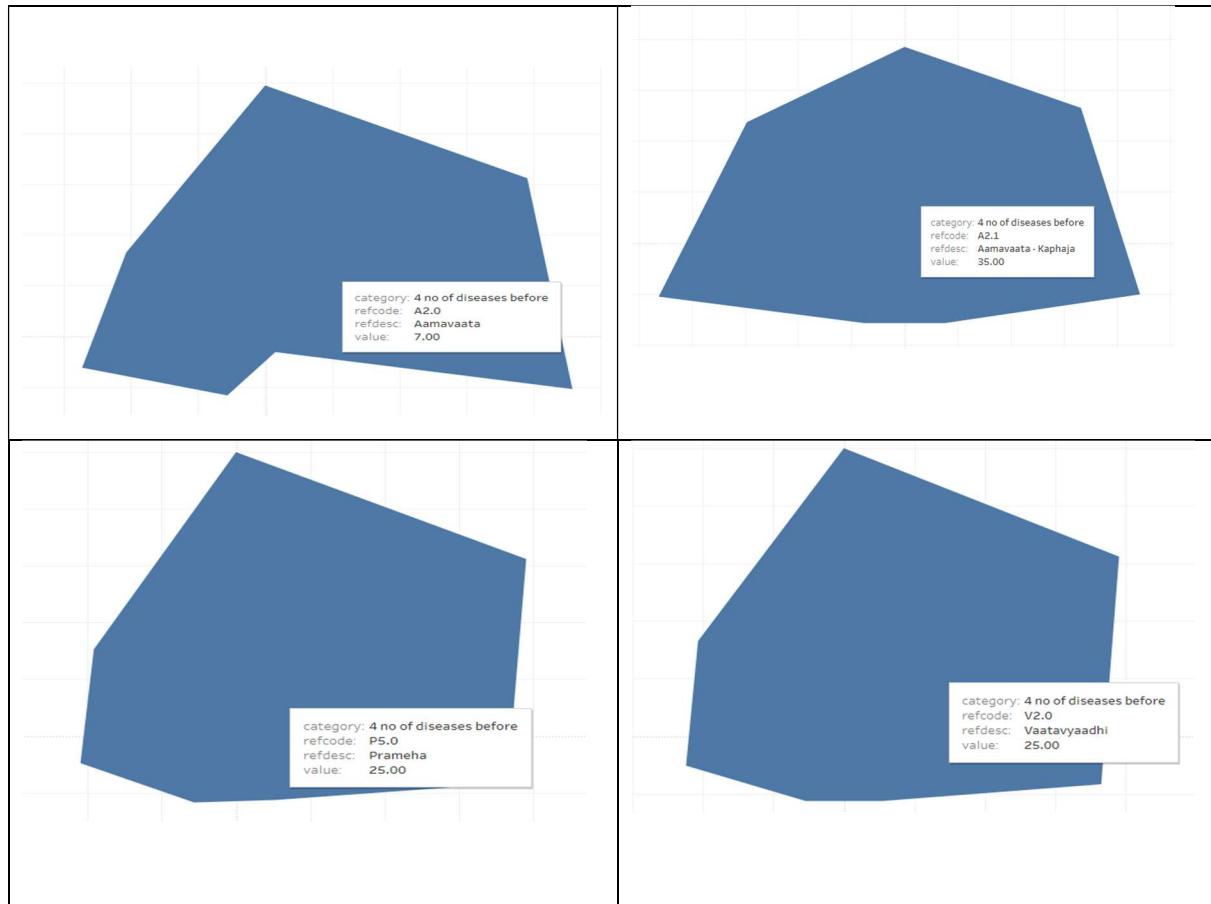
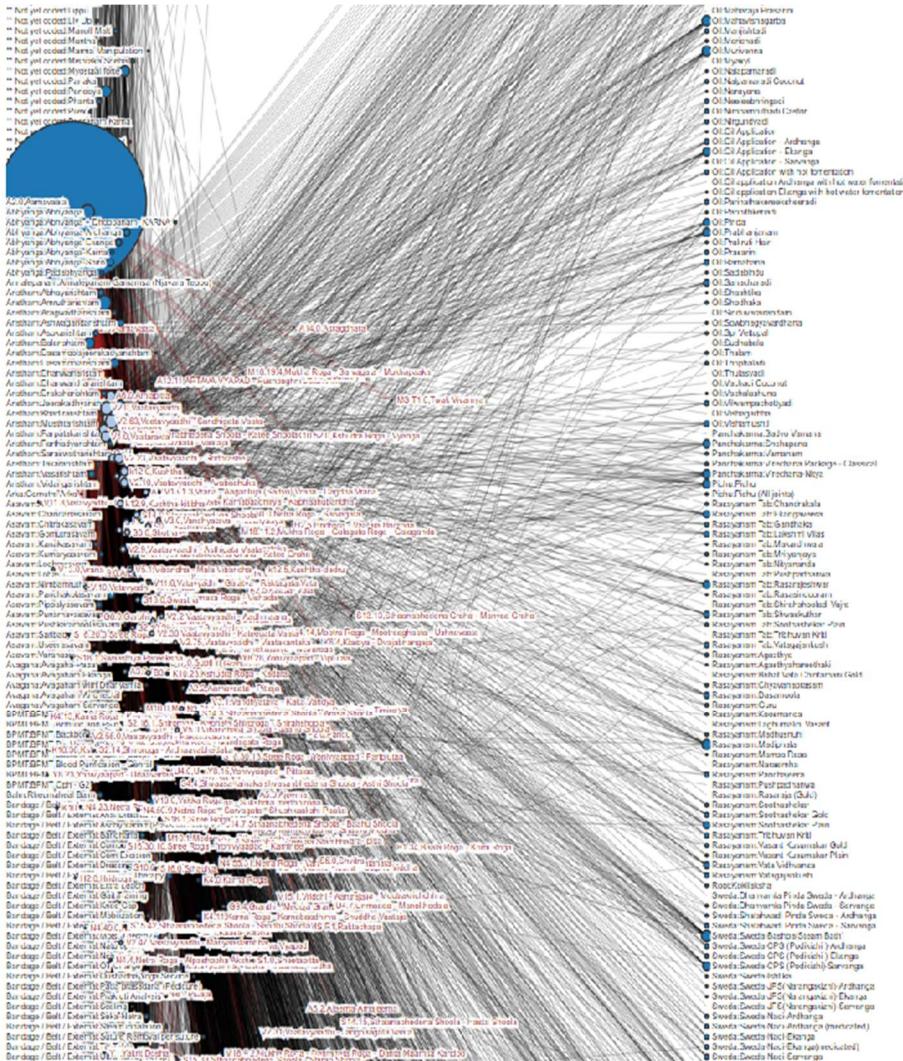


Figure 3-42: Dynamic bubble plot: Example 1: Disease: A6.0: Amavaata



This dynamic bubble plot shows relations between diseases and medicines, the bubble plot size is based on number of unique patients, Data version: 2011 to Oct 2017, Amavaata [[https://coursephd.github.io/nodediagram/A2\\_0/](https://coursephd.github.io/nodediagram/A2_0/)]