

Power BI in Clinical Data exploration, powerful visuals that answer relevant questions

PHUSE Virtual EU Connect 2020

November 9-13, 2020

DEBRUS Roxane

Lead Biostatistician and SAS Programmer

European Medical and Clinical Department, TERUMO EUROPE



Abstract

Nowadays, dynamic data visualization is vital for clinical study conduct and decision making. Power BI is a user-friendly business analytics service by Microsoft which provides interactive visualizations and business intelligence capabilities, offering unique integration options with Python, R, and SQL. Although a leader in BI and Analytics Platforms, its potential in clinical research is currently unexplored. To test it in this capacity and help our clinical and medical teams, we utilized it to design modern-looking, interactive dashboards, with the aim to improve the way we access, review and interact with data. As a result, we developed several tools that facilitate monitoring, analyzing and testing the study data in real time throughout the study, confirming that Power BI is perfectly suited to answer the needs of clinical and medical teams.

Confidentiality Disclosure Statement

The following material contains Terumo confidential and proprietary information. No part of this document may be disclosed, printed, copied or disseminated in any form or by any means, without prior permission from Terumo Europe N.V.

Agenda

■ Introduction :

- Why Data Visualization ?
- Why Power BI ?
- How to develop dashboards in Power BI desktop?

■ Terumo's Power BI Dashboards

- Business Cases Presentation
- Next Steps

■ Conclusion

Introduction

How / Why did we decided to start to use Data Visualization ?

In 2019, Terumo closed one of its largest study: eUltimaster

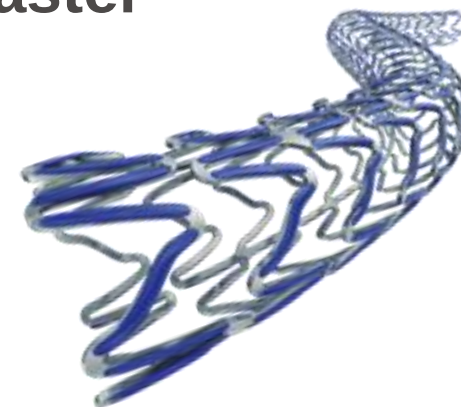
More information available at : <https://www.terumo-europe.com/en-emea/clinicaldata/e-ultimaster-trial>

■ Study Description

- Clinical Field: Cardiovascular Intervention
- Investigational Medical Device : Coronary Stent
- Study Design: Observational study (Registry), single arm, open label, 5 majors timepoints : SCR, BASELINE, PROC, 3MFU, 1YFU
- Primary endpoint: Validate Efficacy and Safety: Composite Endpoint of different SAE up to 1 year after procedure (i.e. TLF = Cardiac Death, TV related to MI and CD-TLR).
- Secondary endpoint: exploratory analysis
- Status: FPI in 2015, DBL 2019, N = 37.198, 378 sites, 50 countries

■ Characteristics

- Global registry, observational study, include all kind of subject, no randomization, subjects treated as per hospital standard of care or PI's preferences.



How / Why did we decided to start to use Data Visualization ?

■ After DBL, Terumo organized a Publication Committee Meeting

■ Top 15 Enrollers were invited to discuss all possible future publications

■ Principal Investigators had many questions !

- They know the patients they included in the study
- However they have no idea what others did
 - How many site / countries participated to the study ?
 - How much did the site / country participate to the study ? Number of subjects ?
First/Last enrollers ? Inclusion duration ? Inclusion speed ?
 - What is the quality of the data collected ? %LFU per site / per country ?
 - What is the typical profile for inclusion ?
Gender / Age distribution ? Lesion Characteristics ?

➔ How to share data overview without having to share full database?

➔ How to share data without compromising data confidentiality?

How / Why did we decided to start to use Data Visualization ?

- Most popular DV platforms in R&D are TIBCO Spotfire, Qlik Sense

- Coming from CRO, I had previous experience with DV (Tableau®)



- Terumo uses Power BI, which is part of Microsoft Office 365

- DV is already in place at Terumo for Finance and IT Departments

- Power BI Desktop is available for any Terumo employee

Introduction



Why Power BI ?

Advantages

- User friendly interface
- Consolidate multiple data source (Excel, CSV, XML, Text, ...)
- Interactive reports and geo-map visualizations
- Integration of R, Python and SQL
- Consistent reporting and analyses
- View reports across multiple platforms and devices
- In constant evolution

Disadvantages

- No direct import of SAS Datasets
(conversion to CSV or Excel needed)

... Or should I say, “Why Not Power BI”?

Magic Quadrant for 2020 for Analytics and Business Intelligence tools:



Introduction

How to develop dashboards in Power BI desktop?

The screenshot displays the Microsoft Power BI Desktop application window. The title bar shows 'Untitled - Power BI Desktop' and the user 'Roxane Debrus'. The ribbon includes 'File', 'Home', 'Insert', 'Modeling', 'View', and 'Help'. The 'Home' ribbon is active, showing sections for 'Clipboard' (Paste, Cut, Copy, Format painter), 'Data' (Get data, Excel, Power BI datasets, SQL Server, Enter data, Recent sources), 'Queries' (Transform data, Refresh data), 'Insert' (New visual, Text box, More visuals), 'Calculations' (New measure, Quick measure), and 'Share' (Publish).

Annotations with green boxes highlight the following components:

- Left Navigation Pane:** Contains icons for Reports, Datasets, and Dataflows. The Reports icon is highlighted.
- Common data sources:** A dropdown menu listing 'Excel', 'Power BI datasets', 'Power BI dataflows', 'SQL Server', 'Analysis Services', 'Text/CSV', 'Web', 'OData feed', and 'Blank query'. The 'More...' option is at the bottom.
- Filters pane:** Includes a search bar, 'Filters on this page' (with 'Add data fields here'), and 'Filters on all pages' (with 'Add data fields here').
- Visualizations pane:** Shows a grid of visualization icons. A dropdown menu is open, listing 'Get more visuals', 'Import a visual from a file', 'Remove a visual', and 'Restore default visuals'. An arrow points from this menu to the 'Get more visuals' link in the bottom right.
- Fields pane:** Includes a search bar and a list of fields: 'ADAE', 'ADSL', 'ACAGE', and 'ACOUNTRY'. Below this, a list of measures is shown: 'AGEGR3N', 'AGEGR4', 'AGEGR4N', and 'AGEU'.

The bottom status bar shows 'Page 1' and a '+' icon for adding new pages. The page number 'PAGE 1 OF 1' is displayed in the bottom left corner.

<https://appsource.microsoft.com/en-us/marketplace/apps?product=power-bi-visuals>

Introduction

How to develop dashboards in Power BI desktop?

The screenshot shows the Power BI Desktop interface. The 'Table tools' ribbon is active, and the 'Calculations' group is highlighted. The 'New measure' button is circled in green. A green arrow points from this button to a callout box titled 'DAX Formula Language'.

The callout box shows the 'Measure tools' ribbon with the 'Name' field set to 'Measure' and the 'Home table' set to 'ADSL'. The 'Structure' tab is active, showing the formula bar with the DAX formula: `1 TOT_USUBJID = DISTINCTCOUNT(ADSL[USUBJID])`.

The background shows a data table with columns: STUDYID, USUBJID, SITEID, SUBJID, BRTHDTC, BRTHY, BRTHDT, BRTHDTF, AGE, AGEU, ACAGE, AGEGR1, AGEGR1N, AGEGR2. The table contains data for various subjects and their ages.

TABLE: ADSL (37,695 rows)

Introduction

How to develop dashboards in Power BI desktop?

The screenshot displays the Power BI Desktop interface. The main ribbon includes tabs for File, Home, and Help. The Home tab is active, showing various tool groups: Clipboard (Paste, Copy), Data (Get data, Excel, Power BI datasets, SQL Server, Enter data, Recent sources), Queries (Transform data, Refresh data), Relationships (Manage relationships), Calculations (New measure, Quick measure, New measure column, New table), Security (Manage roles, View as), and Share (Publish). The Fields pane on the right shows a search bar and a list of fields. The Power Query Editor window is open, showing the Transform tab. The ribbon in the Power Query Editor includes: Home (Close & Apply, New Source, Recent Sources, Enter Data, New Query), Data Sources (Data source settings, Manage Parameters), Query (Refresh Preview, Advanced Editor, Manage), Transform (Choose Columns, Remove Columns, Keep Rows, Remove Rows, Sort, Split Column, Group By, Replace Values, Data Type: Text, Use First Row as Headers, Merge Queries, Append Queries, Combine Files, Text Analytics, Vision, Azure Machine Learning, AI Insights), and a bottom section with various transformation options like Transpose, Reverse Rows, Count Rows, Detect Data Type, Fill, Pivot Column, Convert to List, Split Column, Format, Merge Columns, Statistics, Standard, Scientific, Trigonometry, Rounding, Information, Date, Time, Duration, Structured Column, Run R script, and Run Python script. The 'APPLIED STEPS' pane on the right shows a list of steps: Source, Navigation, Promoted Headers, and Changed Type. The main data view shows a table with columns: STUDYID, USUBJID, SITEID, SUBJID, BRTHDTC, BRTHY, BRTHDT, BRTHDTF, AGE, AGEU, ACAGE, AGEGR1, AGEGR1N, AGEGR2. The table contains 37,695 rows of data.

STUDYID	USUBJID	SITEID	SUBJID	BRTHDTC	BRTHY	BRTHDT	BRTHDTF	AGE	AGEU	ACAGE	AGEGR1	AGEGR1N	AGEGR2
eU-T123E2				1955	1955	Thursday, June 30, 1955	M	59	YEARS	64	< 60 YEARS		1
eU-T123E2				1957	1957	Sunday, June 30, 1957	M	57	YEARS	62	< 60 YEARS		1

TABLE: ADSL (37,695 rows)

Introduction

How to develop dashboards in Power BI desktop?

The screenshot shows the Power BI Desktop interface with the 'Edit relationship' dialog box open. The dialog displays the relationship between the ADSL and ADAE tables. The ADSL table is on the left, and the ADAE table is on the right. The relationship is defined by the STUDYID column in ADSL and the STUDYID column in ADAE. The cardinality is set to 'Many to one (*:1)' and the cross filter direction is 'Both'. The 'Make this relationship active' checkbox is checked.

ADSL Table:

STUDYID	USUBJID	SITEID	ACOUNTRY	RFSTDT	AESEQ	AEGRPID	AE
eU-T123E2				Wednesday, November 5, 2014	1	TBD	
eU-T123E2				Thursday, January 15, 2015	1	TBD	
eU-T123E2				Monday, January 26, 2015	1	TBD	

ADAE Table:

STUDYID	USUBJID	SITEID	SUBJID	BRTHDTC	BRTHY	BRTHDT	BRTHDTF
eU-T123E2				1955	1955	Thursday, June 30, 1955	M
eU-T123E2				1957	1957	Sunday, June 30, 1957	M
eU-T123E2				1942	1942	Tuesday, June 30, 1942	M

Relationship Settings:

- Cardinality: Many to one (*:1)
- Cross filter direction: Both
- ☒ Make this relationship active
- ☐ Assume referential integrity
- ☐ Apply security filter in both directions

Buttons: OK, Cancel

Terumo's Power BI Dashboards

Study Enrollment Dashboard

■ Questions?

- How many site / countries participated to the study ?
- How much did the site / country participate to the study ?
 - Number of subjects ?
 - First/Last enrollers ?
 - Inclusion duration ?
 - Inclusion rate ?
- What is the quality of the data collected ?
 - %LFU per site / per country ?
- What is the typical profile for inclusion ?
 - Gender / Age distribution ?
 - Lesion Characteristics ?



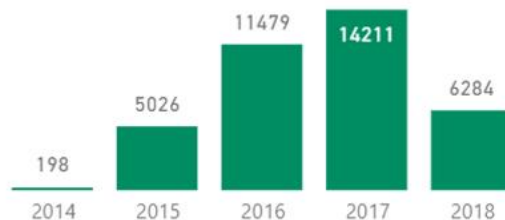


5 Continents - 50 Countries - 378 Sites - 37198 Subjects

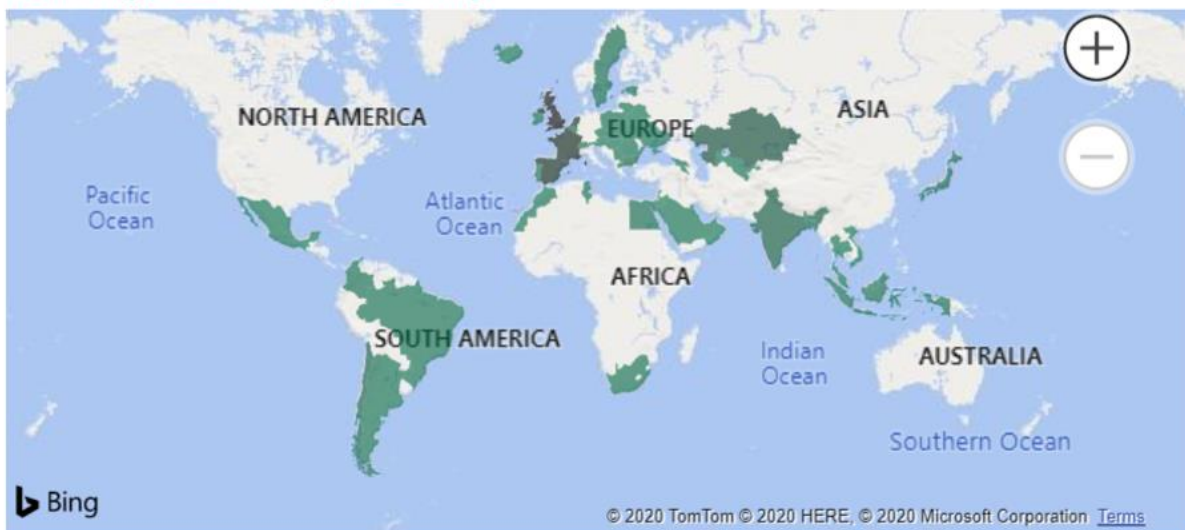
Analysis Population

Baseline Set	
Primary Endpoint Set	Excluded

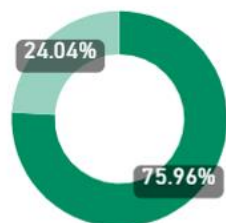
Enrollment per Year



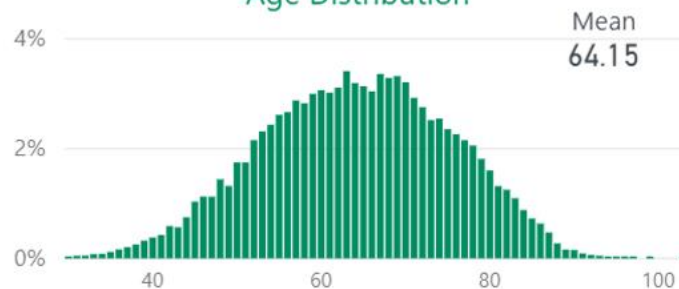
Total Subjects Enrolled per Country



Gender Distribution



Age Distribution



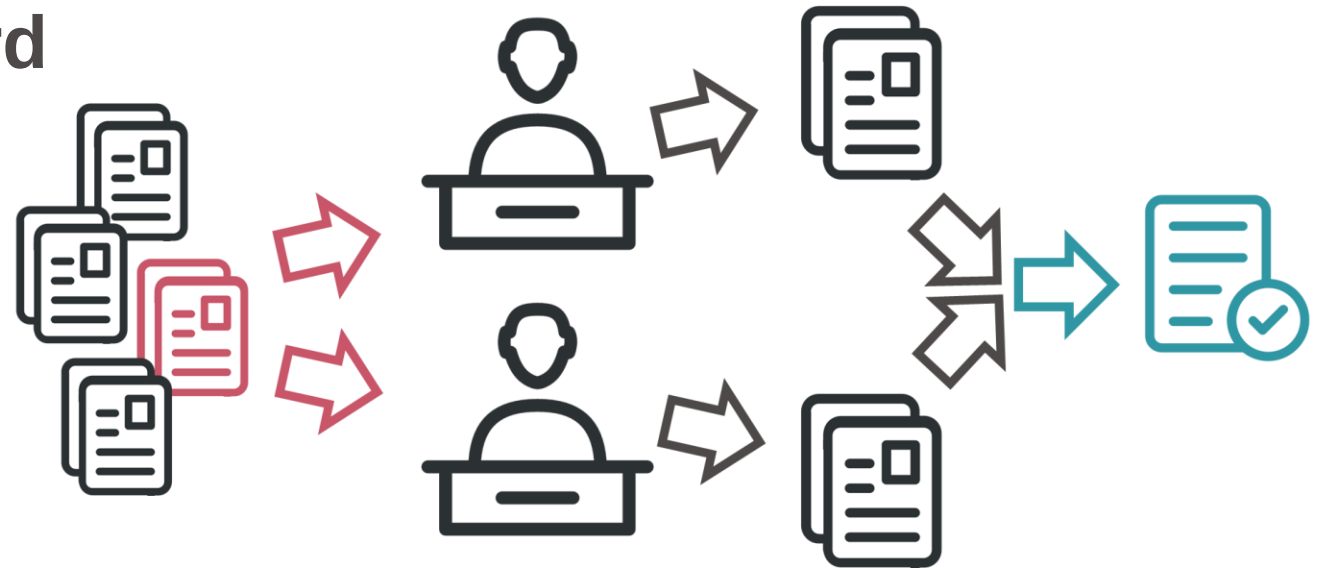
Enrollment Achievement per Continent and by Country

REGION_	N Enrolled	N Site	Mean AGE	%M	FPI	LPI	Incl Dur	Incl Rate	N LFU	% LFU
AFRICA AND MIDDLE EAST	2438	51	59.59	0.80	05-Jan-15	09-May-18	33	73.88	357	14.6
EGYPT	587	13	58.82	0.74	05-Jan-15	06-May-18	29	20.24	154	26.2
ISRAEL	293	4	62.82	0.83	01-Feb-17	01-May-18	11	26.64	6	2.0
JORDAN	28	1	58.25	0.89	10-Dec-16	08-Feb-17	2	14.00		
KUWAIT	50	1	56.18	0.84	10-Apr-17	24-Apr-18	12	4.17	5	10.0
LEBANON	5	2	73.40	0.80	03-Apr-17	10-May-17	1	5.00	5	100.0
MOROCCO	177	3	64.59	0.77	16-Dec-15	09-Apr-18	24	7.38	1	0.6
OMAN	25	1	52.24	0.92	01-Feb-17	28-Apr-17	2	12.50		
SAUDI ARABIA	125	2	55.22	0.61	13-Jan-16	09-May-18	14	8.93	104	83.2
SOUTH AFRICA	295	8	63.93	0.77	11-Nov-15	23-Apr-18	29	10.17	11	3.7
TUNISIA	431	6	64.70	0.77	07-Apr-15	30-Mar-18	33	13.06	6	1.4
UNITED ARAB EMIRATES	422	10	50.11	0.96	17-Sep-15	02-May-18	31	13.61	65	15.4
ASIA	6614	76	60.92	0.76	10-Apr-15	02-May-18	34	194.53	309	4.7
ARMENIA	310	5	60.63	0.83	07-Jun-17	29-Mar-18	9	34.44	68	21.9
BANGLADESH	365	1	54.21	0.91	22-Oct-16	31-Mar-18	17	21.47		
GEORGIA	128	1	63.80	0.71	18-Aug-15	13-Jun-17	22	5.82		
INDIA	1466	14	57.92	0.80	12-Jan-16	01-May-18	27	54.30	17	1.2
INDONESIA	555	8	58.39	0.80	07-Oct-15	02-Jun-17	18	30.83	93	16.8
JAPAN	942	24	69.61	0.77	07-Apr-16	31-Mar-17	11	85.64	66	7.0
KAZAKHSTAN	2319	12	61.33	0.69	10-Apr-15	02-May-18	34	68.21	48	2.1
MALAYSIA	302	6	55.55	0.86	25-Nov-15	26-Apr-18	12	25.17	13	4.3
THAILAND	62	3	62.40	0.81	07-May-16	26-Sep-17	15	4.13	2	3.2
UZBEKISTAN	113	1	60.65	0.74	23-Mar-16	04-Jan-18	22	5.14	2	1.8
VIETNAM	52	1	69.21	0.71	05-Jan-17	17-Apr-18	15	3.47		
EUROPE	25736	222	65.50	0.76	15-Apr-14	02-Jul-18	49	525.22	917	3.6
AUSTRIA	189	3	66.40	0.79	11-Oct-15	30-Apr-18	26	7.27	32	16.9
BELARUS	17	2	61.88	0.82	15-Feb-18	03-May-18	3	5.67		
BELGIUM	613	10	64.92	0.74	15-Jun-15	02-May-18	27	22.70	13	2.1
Total	37198	378	64.15	0.76	15-Apr-14	02-Jul-18	49	759.14	1809	4.9

Safety Monitoring Dashboard

CEC for AESI

Death, Stroke and Revascularization



Questions?

- How many Adverse Event have been adjudicated per site / countries?
- Can we monitor the AESI rate ?
- Do we see difference in AESI rate based on baseline characteristics ?
- How different is data reported in eCRF by Investigators from the Adjudicated Data ?
 - Related to Procedure / Related to Device ?
- Can we try to identify data inconsistency ?
 - Under / Over reporting of AE ? Unexpected rate of a specific AE ?

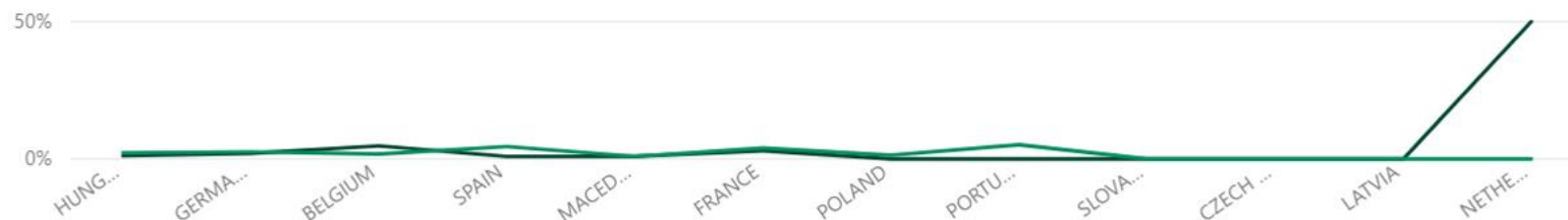




Global Enrollement Achievement



Occurrence Rate of AE



Analysis Subset

N

Y

☐ UP TO 30 DAYS☐ AFTER 30 DAYS

Site Specialization

- ☐ ANGIOLOGIST
- ☐ ANGIOLOGIST - VASC. SURGEON
- ☐ CARDIOLOGIST
- ☐ CARDIOLOGIST - ANGIOLOGIST
- ☐ INTERVENTIONAL NEUROLOGIST
- ☐ INTERVENTIONAL RADIOLOGIST
- ☐ NEURORADIOLOGIST
- ☐ NEUROSURGEON
- ☐ RADIOLOGIST
- ☐ VASCULAR SURGEON

Gender

- ☐
- FEMALE
-
- ☐
- MALE

Status at Baseline

- ☐ ASYMPTOMATIC
☐ SYMPTOMATIC
☐ UNKNOWN

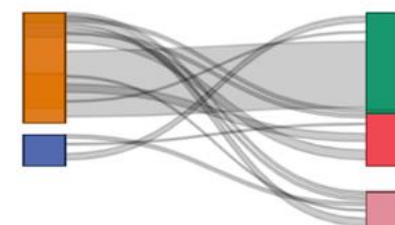
Target Lesion Location

- ☐ BIFURCATION
☐ CCA
☐ ICA
☐ UNKNOWN

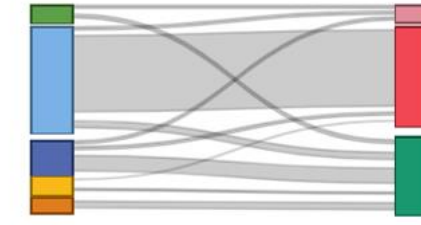
AECAT	DEATH		REVASCULARIZATION		STROKE		Total	
ACOUNTRY	N Subject	N AE	N Subject	N AE	N Subject	N AE	N Subject	N AE
BELGIUM	8	8	16	16	3	3	24	27
FRANCE	3	3	4	4	4	4	10	11
GERMANY	4	4	12	12	5	5	16	21
HUNGARY	4	4	3	3	8	9	8	16
MACEDONIA	1	1			1	1	1	2
NETHERLANDS	1	1					1	1
POLAND					1	1	1	1
PORTUGAL					2	2	2	2
SPAIN	1	1	4	4	5	5	7	10
Total	22	22	39	39	29	30	70	91



Relationship to Device



Relationship to Procedure



Detailed Listing of Adjudicated Adverse Event

COUNTRY	SITE ID	USUBJID	CATEGORY	SUBCATEGORY	RELATIVE START DATE	RELATIONSHIP TO PROCEDURE	RELATIONSHIP TO DEVICE	LOCATION	TYPE	CIRCUMSTANCE
BELGIUM	1	[REDACTED]	DEATH		200	NO	NO			CARDIAC RELATED
BELGIUM	1	[REDACTED]	DEATH		57	NO	IMPOSSIBLE TO DETERMINE			UNKNOWN
BELGIUM	2	[REDACTED]	DEATH		92	NO	NO			OTHER: LUNG CANCER
BELGIUM	2	[REDACTED]	DEATH		85	NO	NO			OTHER: LUNG CANCER
BELGIUM	2	[REDACTED]	REVASCULARIZATION	NON-TARGET VESSEL	205	NO	NO		OTHER: CONTRALATERAL ICA STENOSIS	
BELGIUM	2	[REDACTED]	REVASCULARIZATION	NON-TARGET	63	NO	NO		OTHER:	

Lesion Characteristics and Impact on outcome

■ Questions?

- What are the lesion characteristics of the treated subjects ?
- How does these characteristics impact the studied outcome ?





Baseline Set

Primary Endpoint Set

Excluded

35388 Subjects - 47442 Lesions

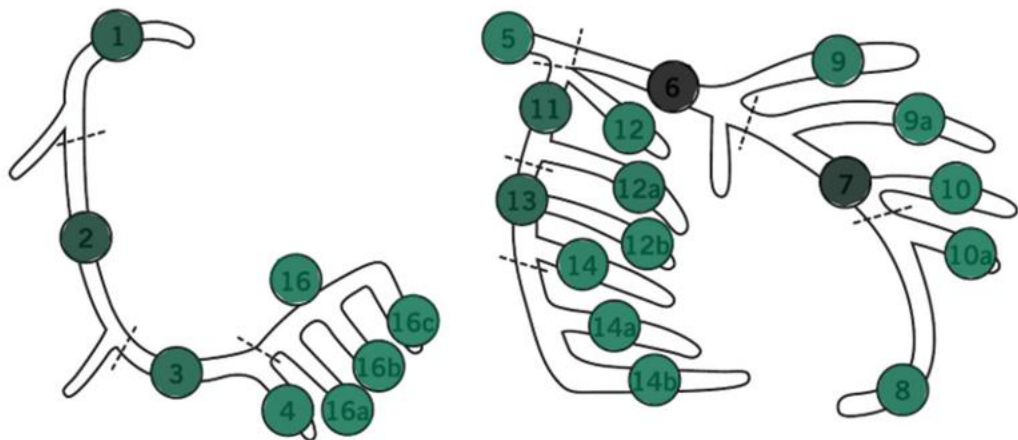
INDEX PROCEDURE

STAGED PROCEDURE #1

STAGED PROCEDURE #2

STAGED PROCEDURE #3

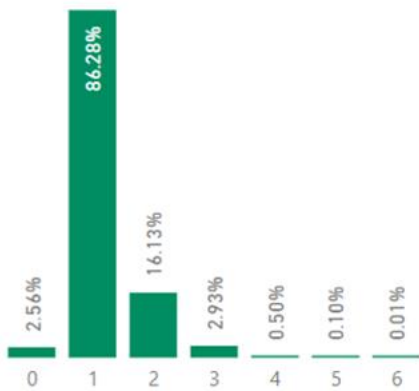
Lesion Description per Segment



Lesion Type Description



Number of Stent Implanted



AHA Classification



Balloon Pre-dilatation



Balloon Post-Dilatation



1 Year Outcome by Treated Vessel and by Treated Segment

SEGLOC	N Lesions	TLF1YN	TLF1Y%	DTHCU 1YN	DTHCU 1Y%	MITVIM 1YN	MITVIM 1Y%	RVTLCD YN	RVTLCD 1Y%
LAD									
6	20000	801	2.80%	241	1.2%	170	0.85%	819	1.83%
7	8010	100	1.25%	100	1.25%	100	1.25%	100	1.25%
9	8017	200	2.50%	112	1.4%	78	0.97%	112	1.40%
8	1400	50	3.57%	10	0.71%	10	0.71%	20	1.43%
91	140	40	28.57%	10	7.14%	11	7.86%	22	15.71%
10	140	4	2.86%	3	2.14%	1	0.71%	3	2.14%
101	14	1	7.14%	1	7.14%	1	7.14%	4	28.57%
102	10	0	0.00%	0	0.00%	0	0.00%	0	0.00%
RCA									
2	14074	810	5.75%	114	0.8%	100	0.71%	100	0.71%
1	14072	100	0.71%	10	0.07%	10	0.07%	10	0.07%
3	4017	112	2.79%	40	0.99%	30	0.75%	10	0.25%
4	2010	70	3.48%	20	0.99%	20	0.99%	20	0.99%
16	101	10	9.9%	10	9.9%	10	9.9%	10	9.9%
161	100	10	10%	10	10%	10	10%	10	10%
163	41	0	0%	0	0%	0	0%	0	0%
162	20	0	0%	0	0%	0	0%	0	0%
LCX									
11	10001	204	2.04%	100	1.0%	80	0.8%	170	1.7%
13	1400	100	7.14%	40	2.86%	40	2.86%	40	2.86%
121	1100	10	0.91%	10	0.91%	10	0.91%	10	0.91%
12	1100	40	3.64%	10	0.91%	10	0.91%	20	1.82%
14	100	10	10%	10	10%	10	10%	10	10%
122	200	10	5%	10	5%	10	5%	10	5%
15	100	10	10%	10	10%	10	10%	10	10%
141	10	0	0%	0	0%	0	0%	0	0%
142	10	1	10%	1	10%	1	10%	1	10%
LM									
5	1100	100	9.09%	10	0.91%	10	0.91%	10	0.91%
Total	47442	1100	2.32%	400	0.84%	310	0.65%	800	1.68%

Complex PCI Dashboard

■ Complex PCI subjects defined by 6 criteria

- Bifurcation lesion treated with a 2-Stents technique
- Chronic Total Occlusion
- ≥ 3 Lesions Treated
- Multi Vessel Treated
- ≥ 3 Stents Implanted
- Total Stent Length ≥ 60 mm

■ Questions ?

- Distribution per criteria ? N ? Overlap ?
- Impact of each criteria on specific outcomes ?
- Most uses Implanted Stents sizes ?



STUDY DASHBOARD - COMPLEX PCI

Database Extraction Date : 09-DEC-2019

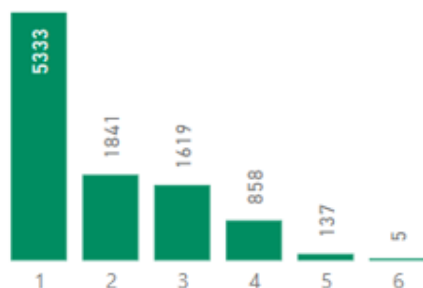
Ultimaster™
Sirolimus Eluting Coronary Stent
System



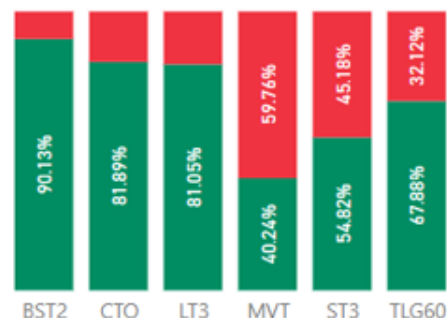
9793 Subjects

Baseline Set	
Primary Endpoint Set	Excluded
COMPLEX	NON COMPLEX

Number of Complex Criteria



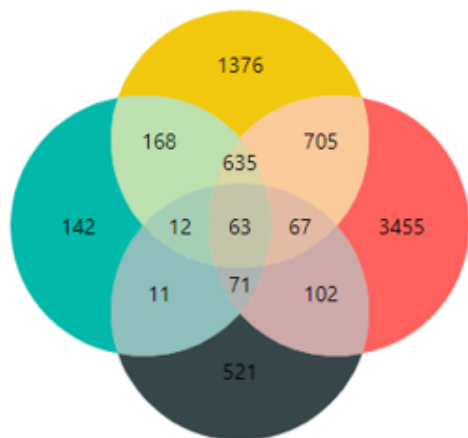
Criteria Distribution



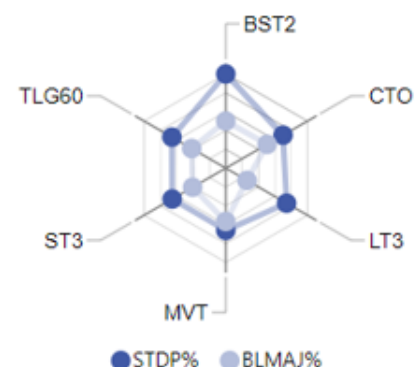
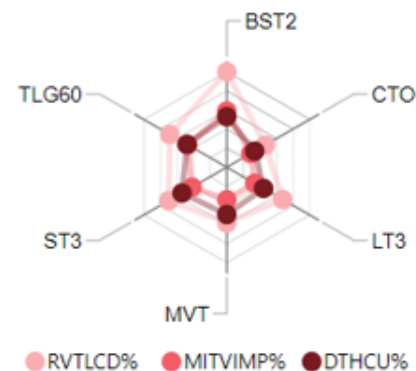
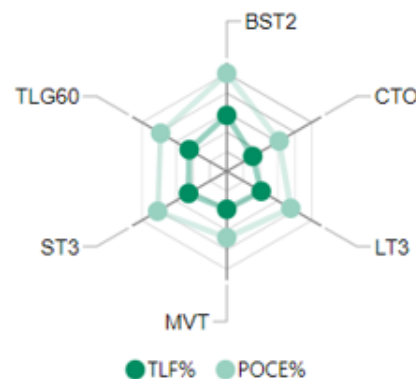
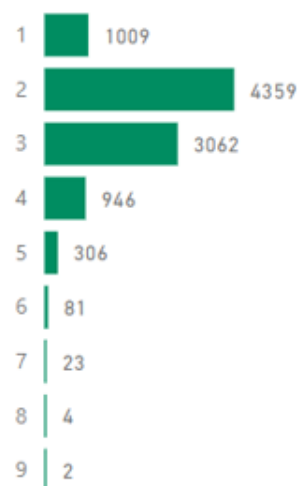
Distribution of Subjects among the Complex Criteria

Legend

- LT3: 1856, 19%
- BST2: 967, 10%
- MVT: 5852, 60%
- TLG60: 3146, 32%



Number of Stent Implanted

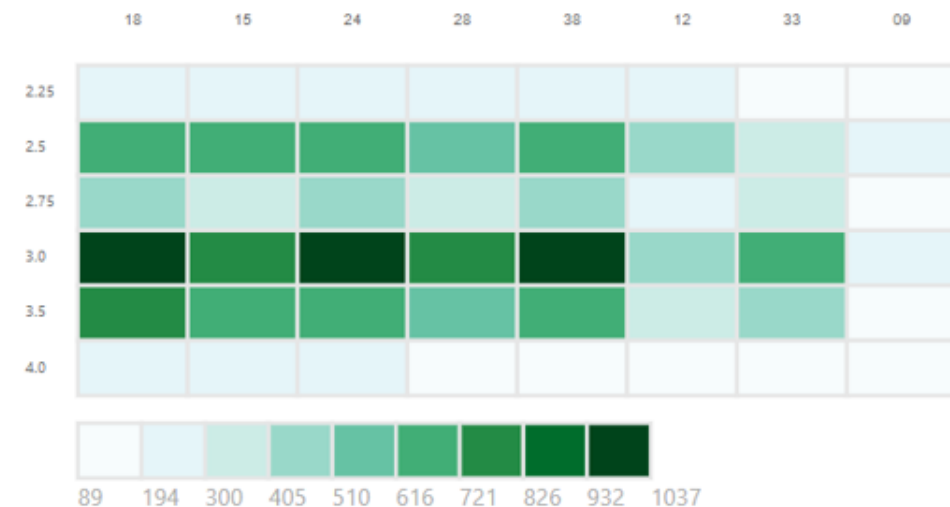


1 Year Outcome

CPXPCI	N Subjects	DTHCU%	MITVIMP%	RVTLCD%	TLF%	POCE%	STDP%	BLMAJ%
N	25596	1.16	0.78	1.49	2.84	5.99	0.58	0.49
Y	9793	1.60	1.19	2.14	4.17	7.90	0.92	0.78
Total	35389	1.29	0.89	1.67	3.21	6.52	0.67	0.57

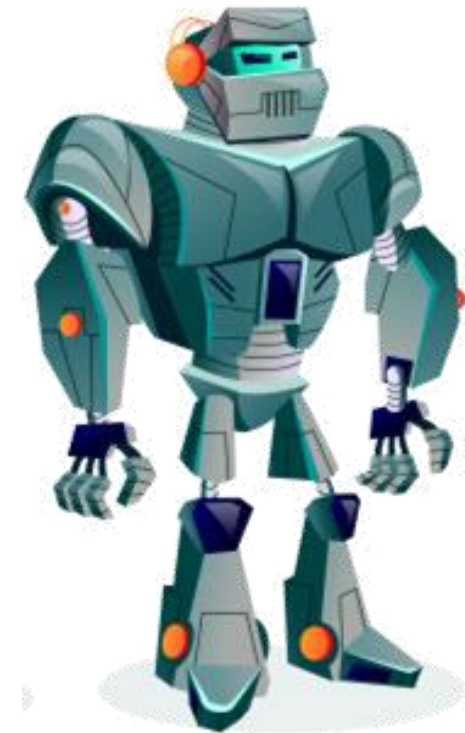
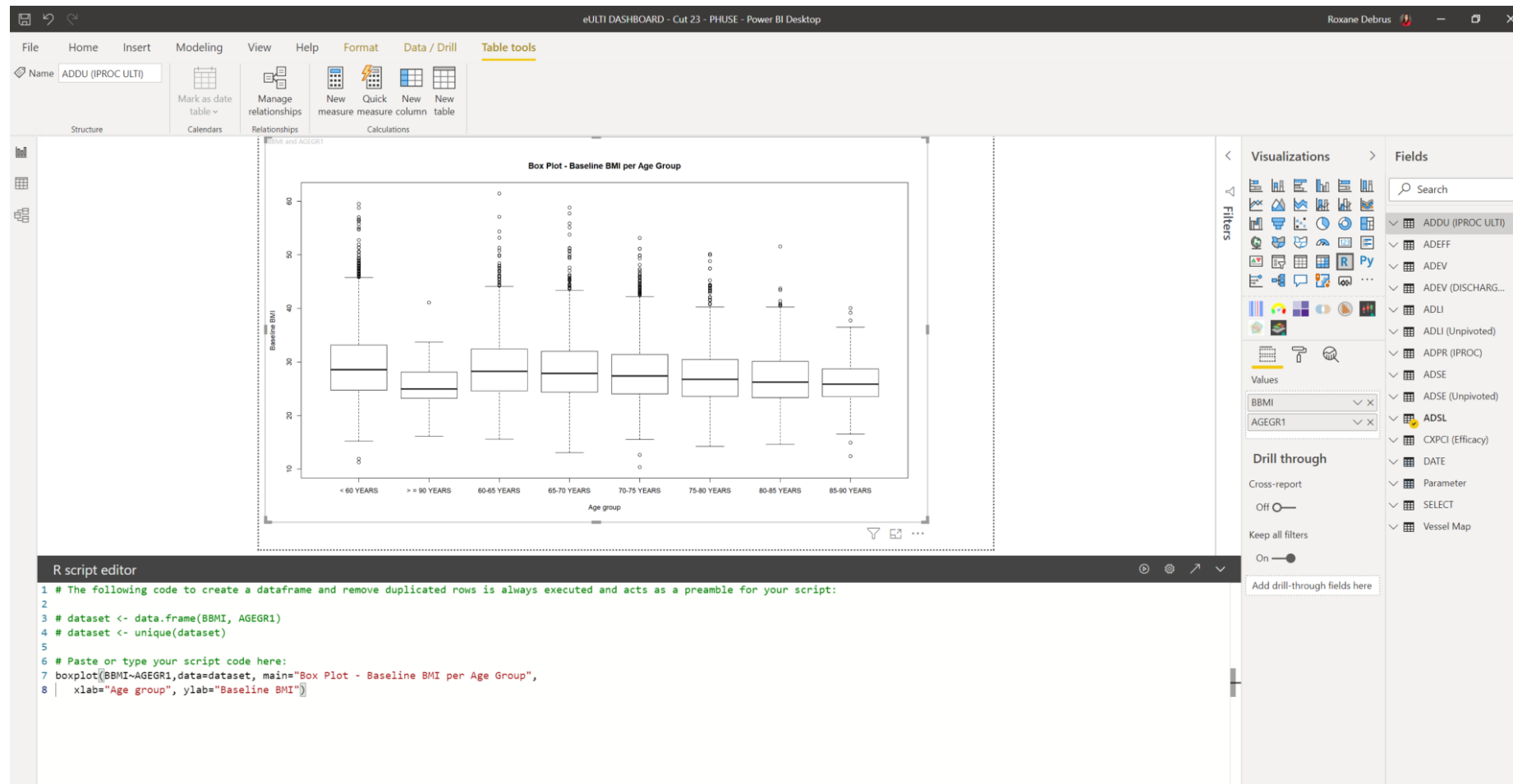
CXPCAT	N Subjects	DTHCU%	MITVIMP%	RVTLCD%	TLF%	POCE%	STDP%	BLMAJ%
BST2	967	1.86	2.07	3.52	6.20	10.86	1.45	0.72
CTO	1774	1.18	1.01	1.63	3.33	6.71	1.01	0.73
LT3	1856	1.56	1.19	2.37	4.42	8.19	1.08	0.38
MVT	5852	1.74	1.20	2.05	4.20	7.35	0.96	0.82
ST3	4424	1.90	1.47	2.46	4.88	8.84	0.95	0.59
TLG60	3146	1.68	1.65	2.42	4.80	8.46	0.95	0.60
Total								

Stent Implanted : Diameter (mm) vs Length (mm)



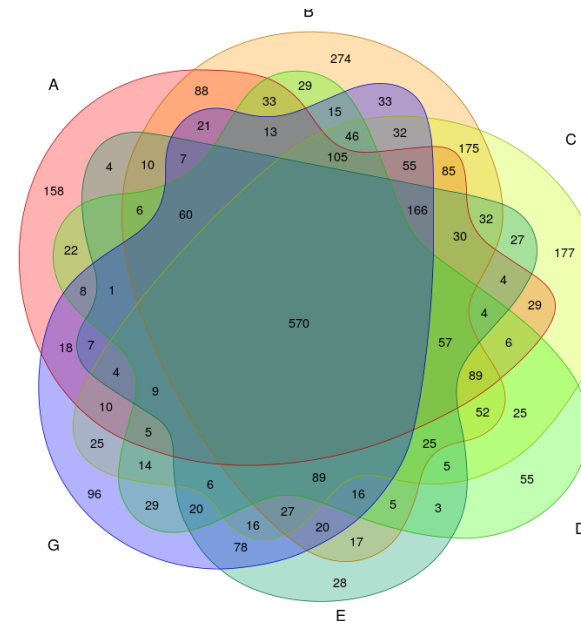
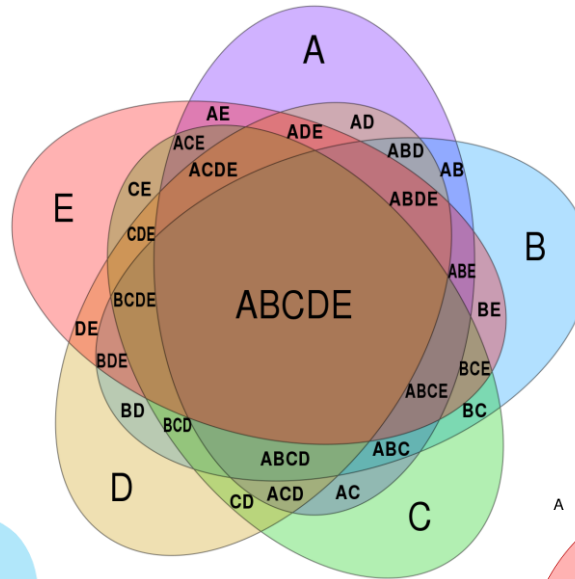
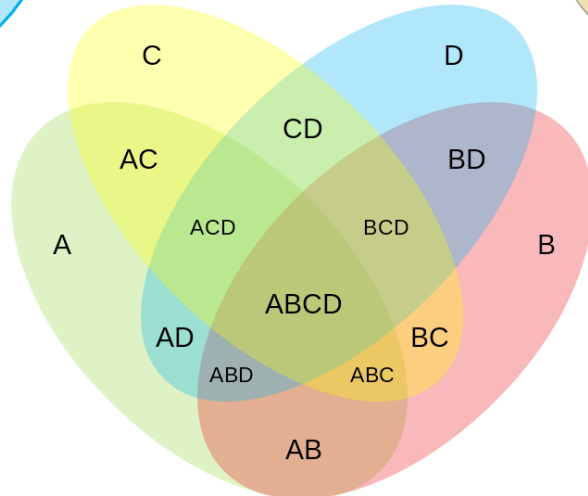
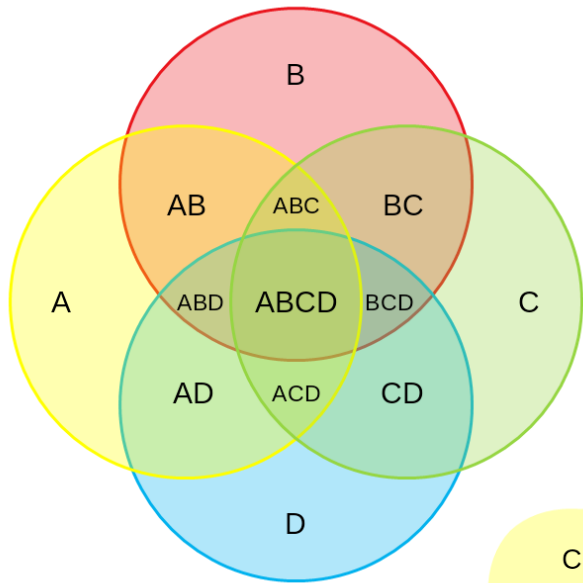
What's Next ?

Integrate user defined R script



What's Next ?

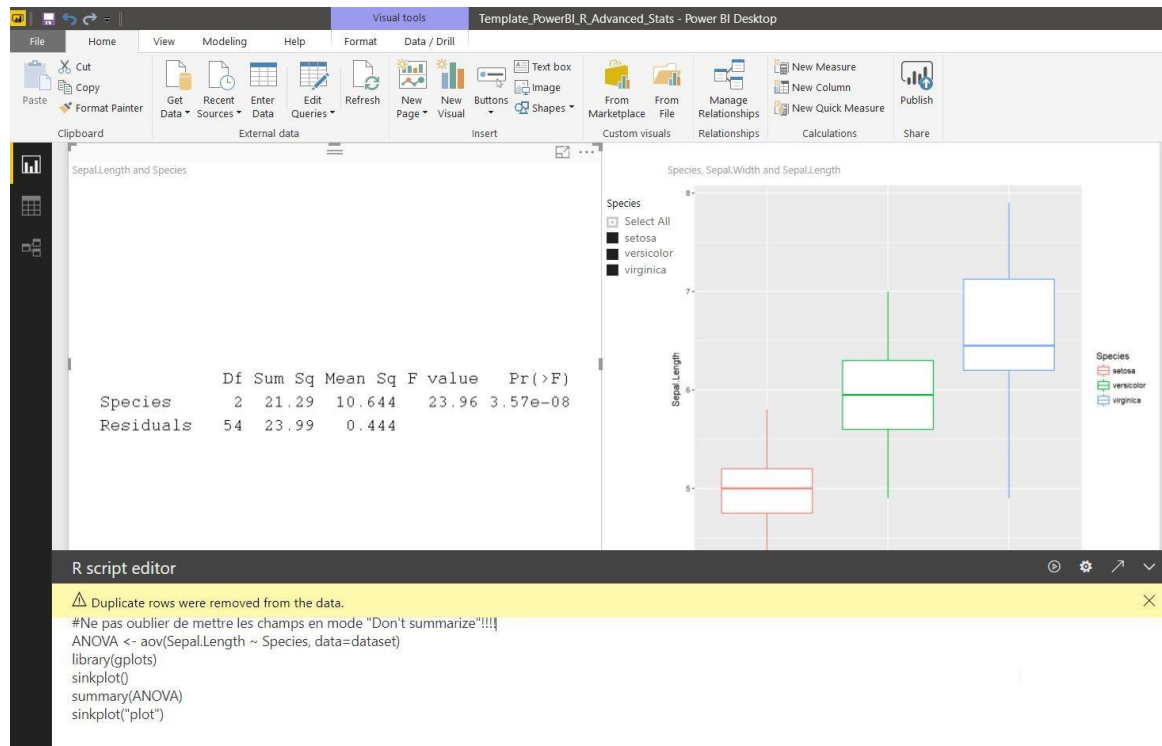
- Integrate user defined R script – to further develop visuals



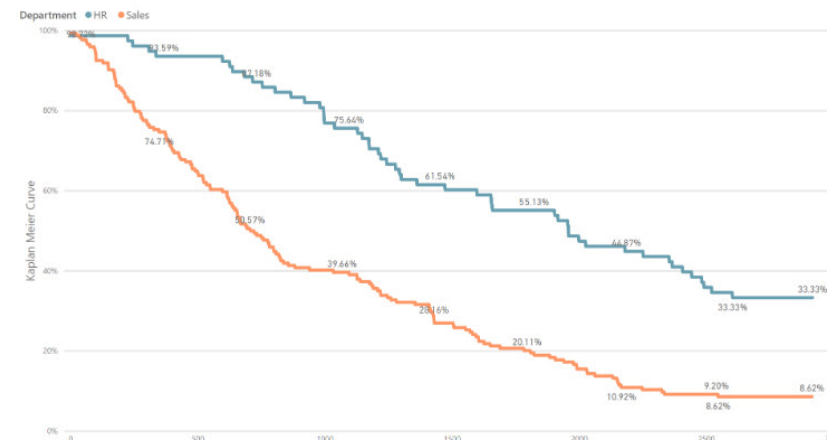
What's Next ?

Integrate user defined R script – to integrate statistical output

ANOVA

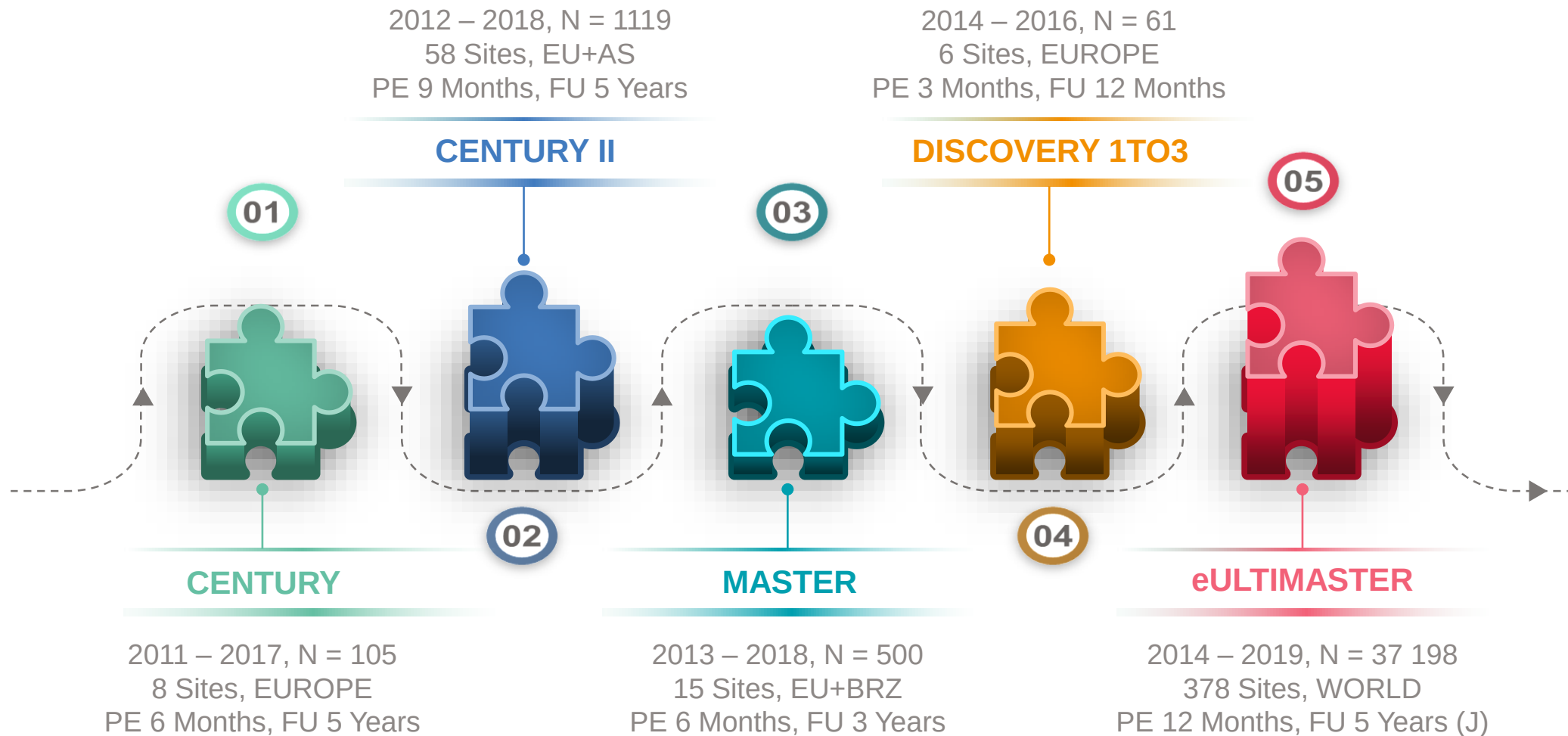


KAPLAN MEIER CURVES



What's Next ?

■ Pooling of different studies to get more global picture



What's Next ?

- Pooling of different studies to get more global picture



Conclusion

Conclusion

■ How to start ?

■ Define your user story !

- Try to identify what you miss
- Define a question you want to answer
- Define your audience and their expectations

■ Define the data to be used (raw vs standardized)

■ Start small and increase complexity progressively

■ Do not overcomplicate things

■ Do not link data randomly, start from a medical question, does “X impact Y ?”



Thank you!