

The Human in the Loop: Accelerating AI/ML Efficiencies for Clinical Data

Achilleas Zaras

Senior Manager of Solution Consulting
eClinical Solutions

PHUSE EU Connect 2023



Agenda

Terminology Refresher

Industry Drivers

AI Hype Cycle

Demystifying the Hype

HILT Overview

HILT – Development

Technology Foundation

AI Enabled Use Cases

Case Study

Key Takeaways

Q & A



Terminology Refresher



Artificial Intelligence: is the capability of a computer system to mimic human cognitive functions such as learning and problem-solving. Through AI, a computer system uses math and logic to simulate the reasoning that people use to learn from new information and make decisions.

Machine Learning: is when we teach computers to extract patterns from collected data and apply them to new tasks that they may not have completed before.



Driving the evolution

INDUSTRY DRIVERS

- ✓ Trial Complexity
- ✓ The 5 V's
- ✓ Risk Based Strategies
- ✓ Decentralized Trials

As a cornerstone of the end-to-end drug development lifecycle, Clinical Data Management teams must acknowledge what is driving the overarching need to evolve and leverage AI and ML technologies to digitally transform data management.

From Clinical Data Manager (CDM) to Clinical Data Sciences (CDS)

Clinical Data Management

Storing data in
a static form

Manual
Review

Logical Checks

100% SDV

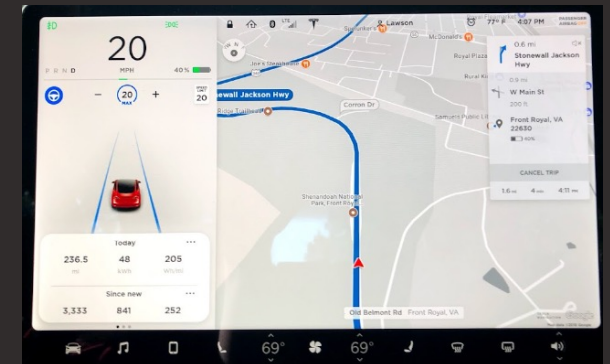
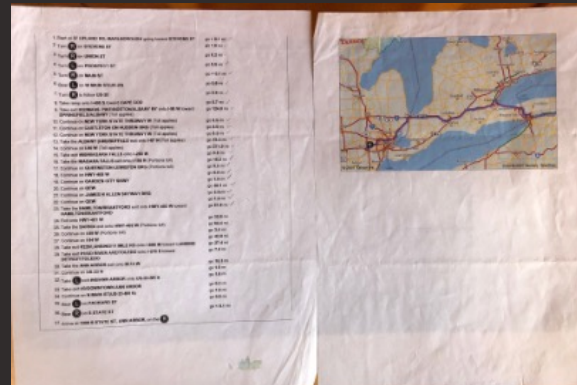
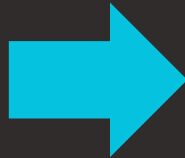
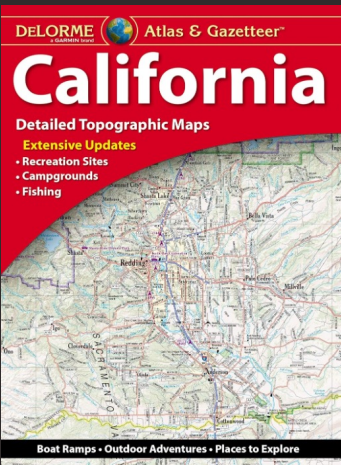
Clinical Data Sciences

Real Time
Insights

Predictive
Analytics

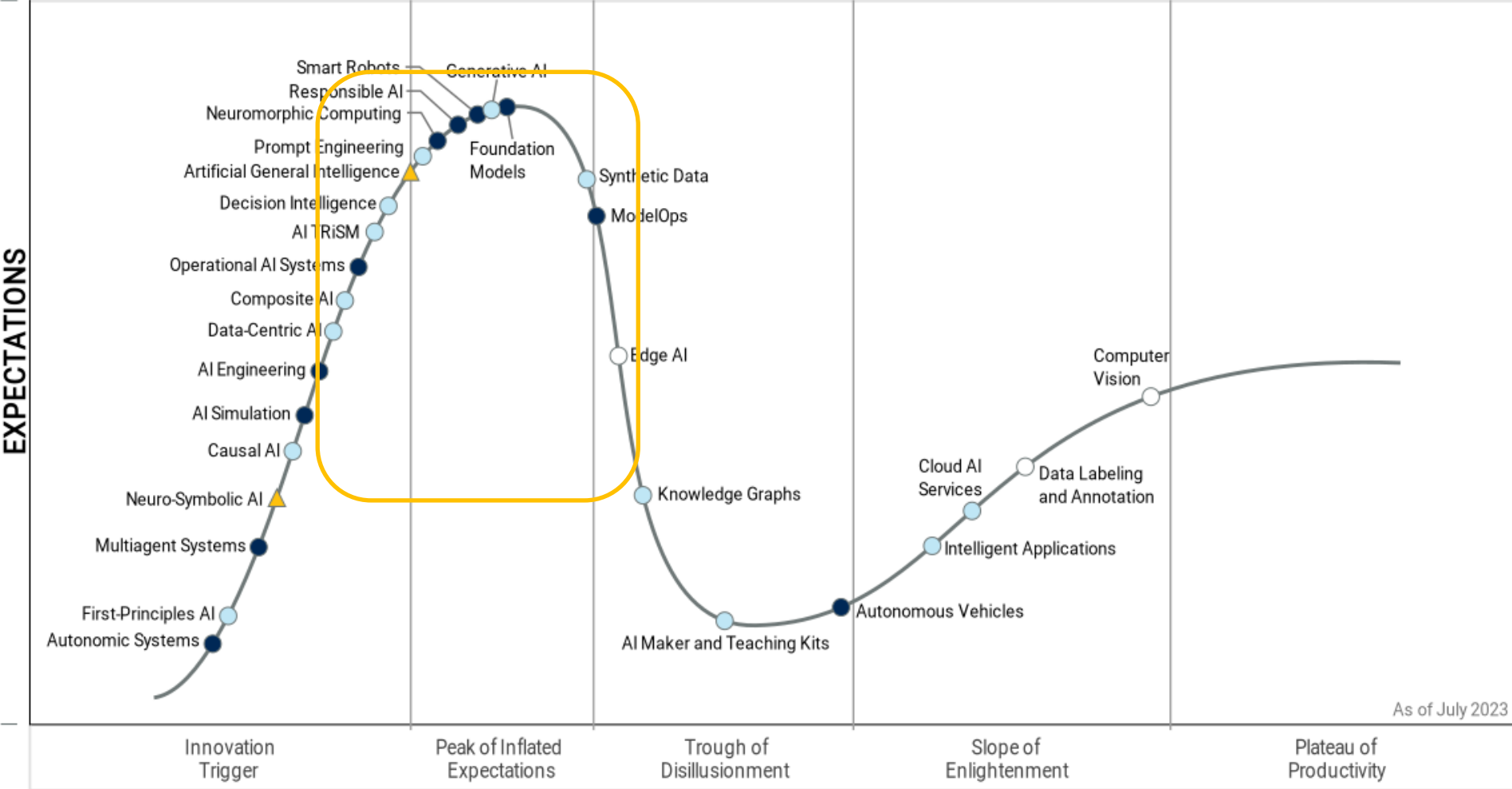
Anomaly
Detection

Meaning &
Value



Hype to Reality

Hype Cycle for Artificial Intelligence, 2023

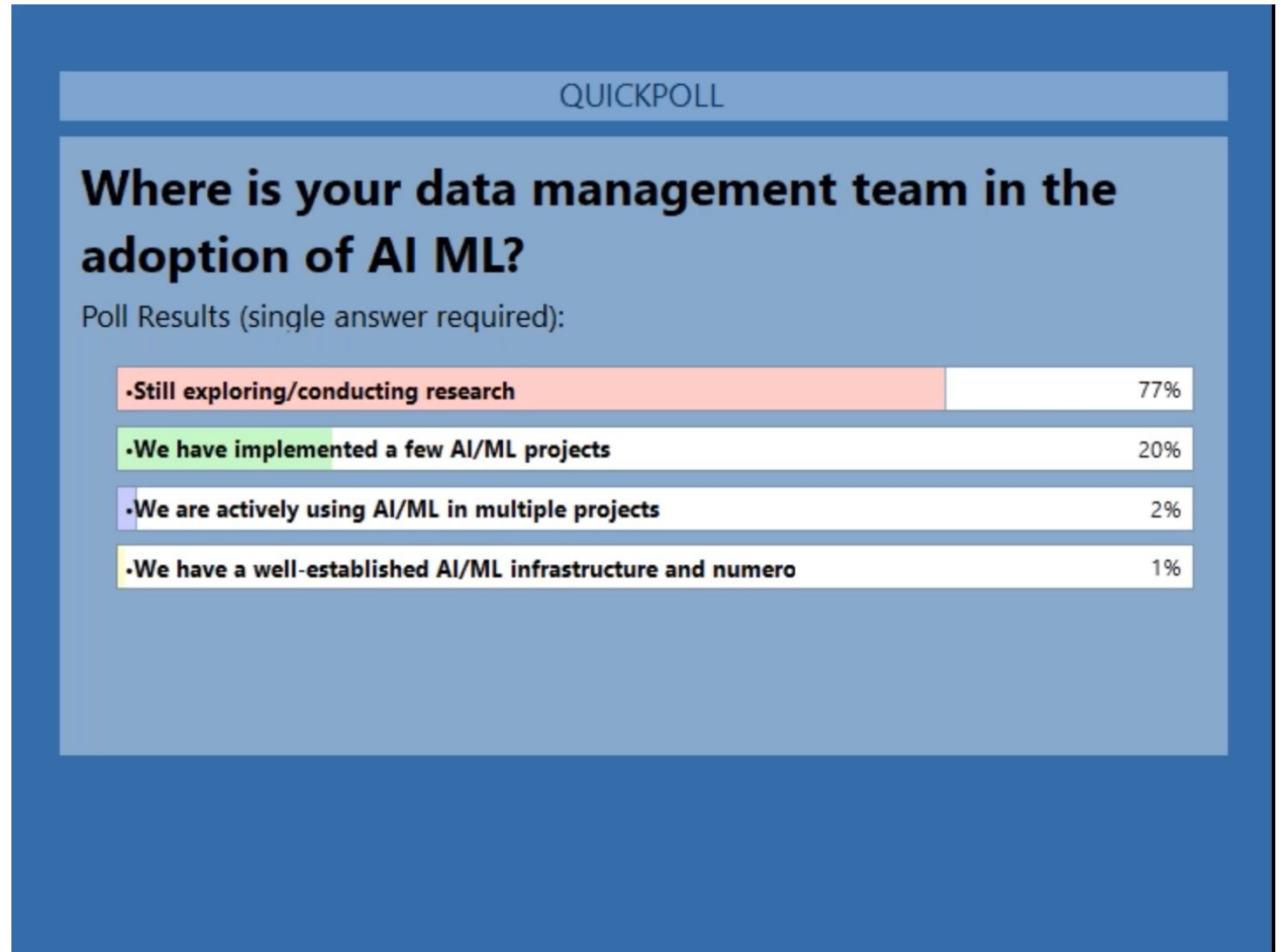


Plateau will be reached: ○ <2 yrs. ● 2-5 yrs. ● 5-10 yrs. ▲ >10 yrs. ✗ Obsolete before plateau



Recent Poll Question

- Where is your data management team in the adoption of AI ML?



Demystifying the Hype

Traditional Data Review Methods

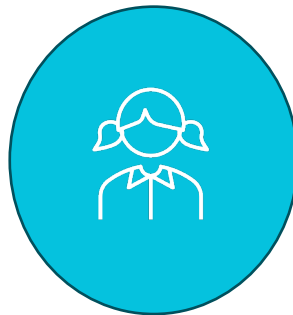
- Time Consuming
- Error Prone
- Low accuracy
- Resource Intensive

Manual Reviews

DM Initiated Queries

Decision Making

Action



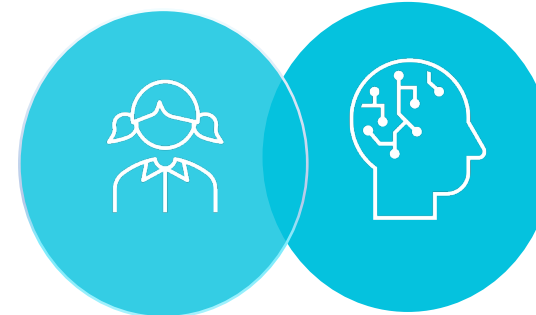
AI Augmented Data Review Methods

AI ML Review
of Focused
Output

Automation of
Proposed
Query Text

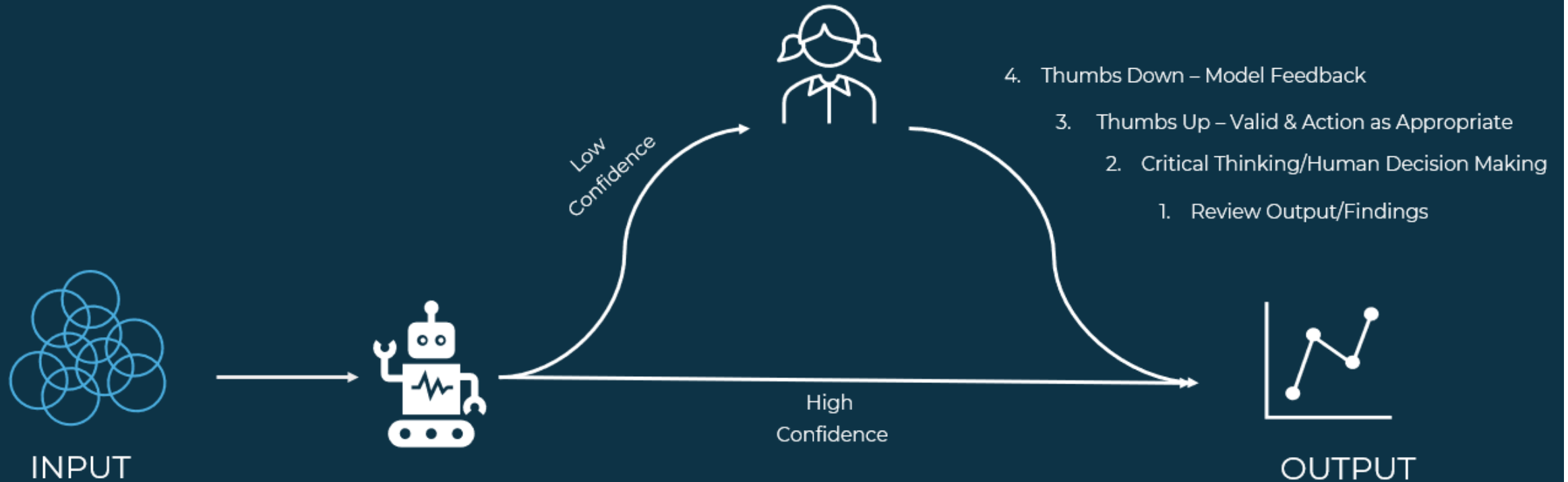
Human in the
Loop

Action



- Time Saving
- Higher Accuracy
- Higher efficiency and productivity

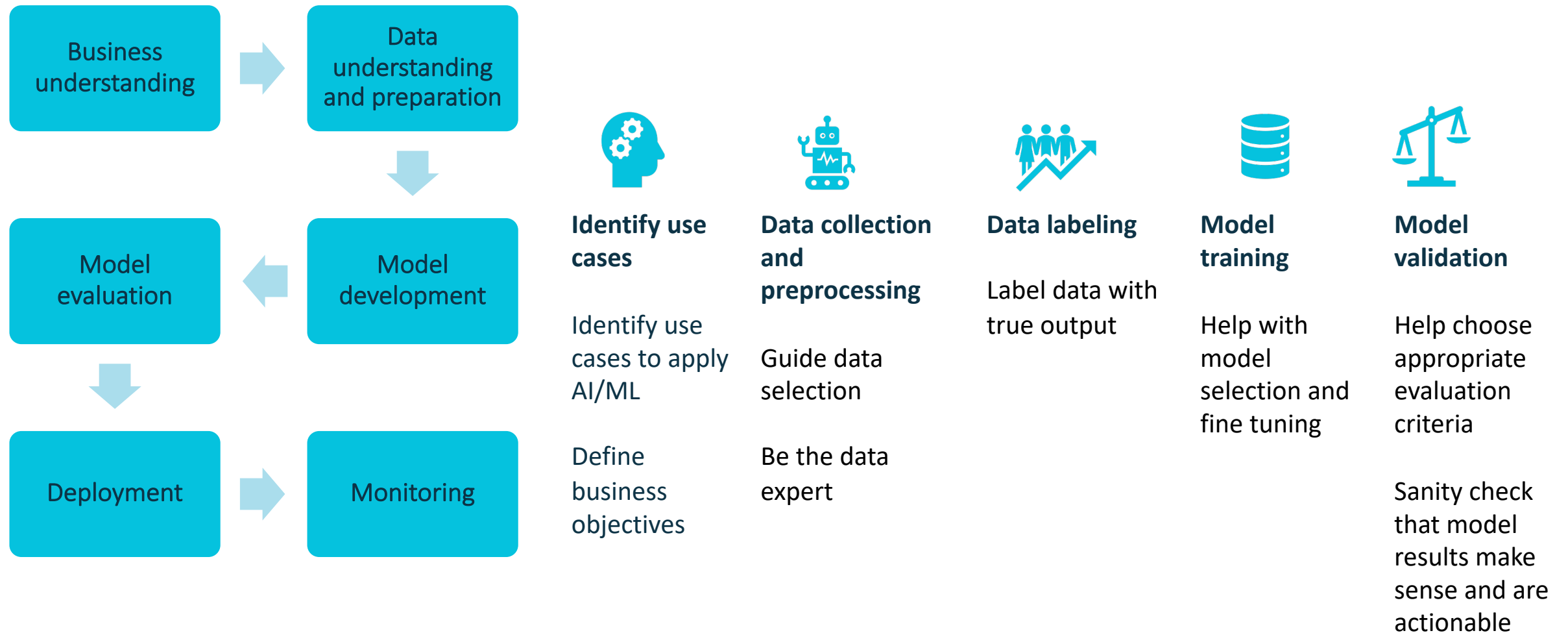
Human in the Loop (HITL)



Data Manager as the Human in the Loop –
collaborating to bring AI ML from ideation to reality



DM as HITL - Development



Example: CM-Indication Appropriateness

CMs Containing Anomalies 5			
CMDECOD	# of Analyzed Subjects	# Analyzed Records	# of Anomalous Sub
Alprazolam	20	42	3
Aprepitant	18	47	4
Calcium	34	50	2
Dexamethasone	15	45	2
Fentanyl	18	47	4

Anomalous Records for Dexamethasone 3							
		MDR	Subject ID	Record ID	CMINDC	Site ID	Reason
☒	☐		CLINTEK-012-007-002	2512	Hydration	87-007	The med
☐	☐		CLINTEK-012-007-002	2624	Prophylaxis	87-007	The med
☐	☐		CLINTEK-012-009-003	9854	Hypotension	87-009	The med

Flag CM records where the medication given for an indication may not be appropriate and DM should review.

HITL:

- **Identify use cases:** Need to reduce workload and improve consistency and accuracy
- **Data cleaning and preprocessing:** Identify study-dependent rules to pull domains where medications and indications are located
- **Data labeling:** Label CM-indication data as appropriate/not appropriate. Match CM with open-source data.
- **Model validation:** Assess performance for go/no-go decision. Select operating point to balance performance.
- **Deployment:** Guide user interface.

Recent Poll Question

- As a Data Manager, what do you consider to be the most significant challenges in adopting Human in the Loop (HITL) AI?

QUICKPOLL

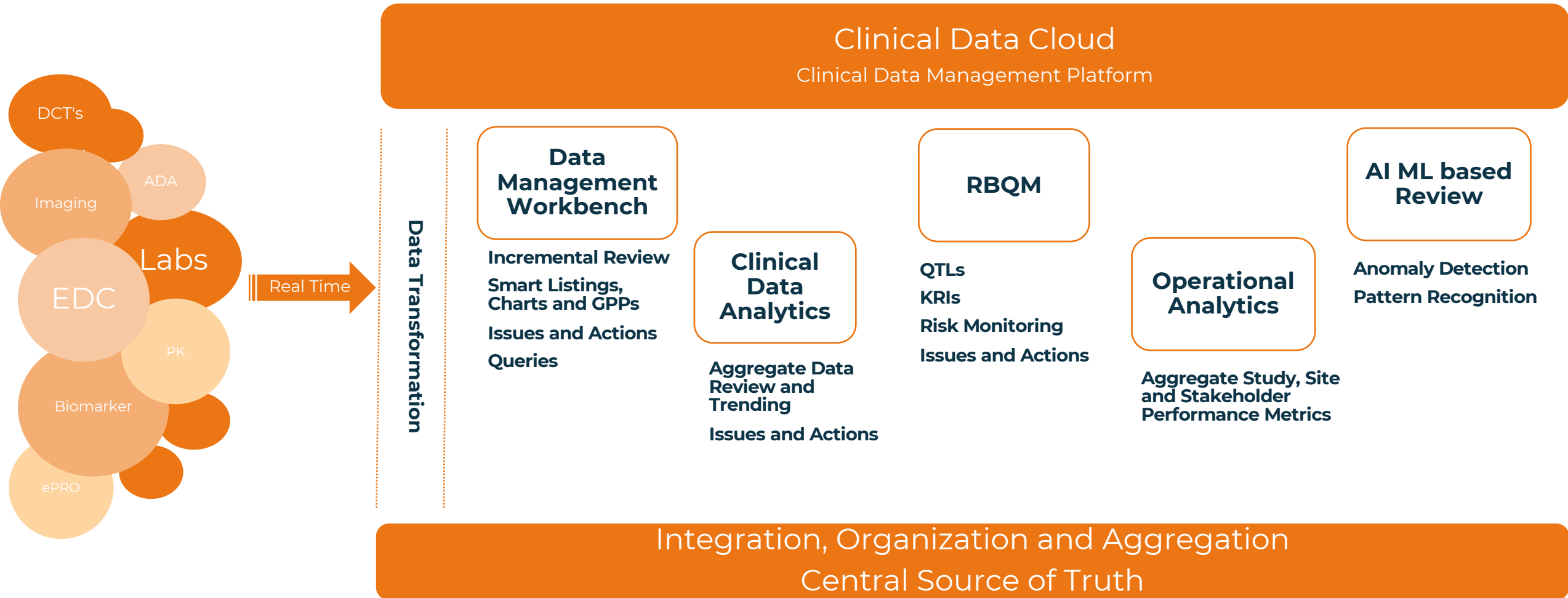
As a Data Manager, what do you consider to be the most significant challenges in adopting Human in the Loop (HITL) AI?

Poll Results (single answer required):

•Navigating and mastering the new technology	34%
•Finding the time to actively engage with AI systems	13%
•Developing confidence in AI systems' abilities and decisio	33%
•Concerns about how personal data is used and protected	9%
•Fears of job displacement or changes in job roles	10%

AI Enabled Use Case Review

Foundation for AI ML



Case Study

The background of the slide is a close-up, high-angle photograph of a microchip or integrated circuit. The chip's surface is covered in a dense, intricate pattern of fine lines, which are the circuit traces, and various rectangular and square components, likely transistors or memory cells. The lighting is soft, highlighting the texture and complexity of the silicon surface. A dark teal horizontal bar is superimposed over the middle of the image, containing the title text in white.

Case Study and Key Takeaways

Case Study: Traditional Data Review vs AI Augmented Data Review

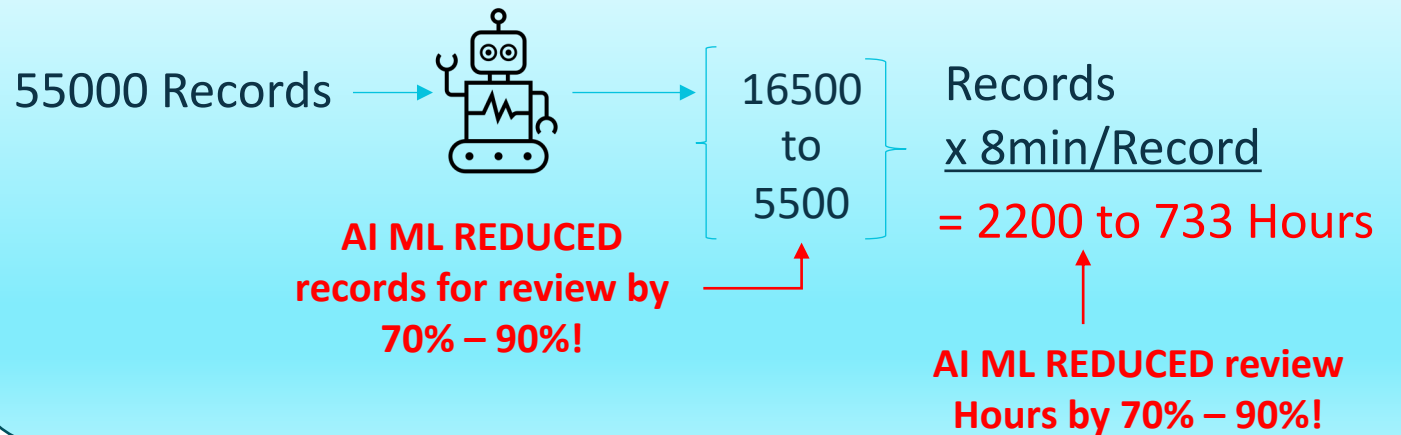
Use Case: Traditional Data Review vs AI ML HITL Review

Parameters: Use of 4 AI ML Models

Traditional Data Review

55000 Records
 $\times \text{8 min/Record}$
= 7333 Hours

AI ML HITL Review



Key Takeaways

1

AI ML Reduces Effort

AI ML flags data anomalies and inconsistencies and allowing data managers to focus their efforts on resolving specific issues vs manually reviewing data.

2

Human Expertise is Essential

AI ML can only succeed with human professionals' supervision, expertise, and contribution – enter the data manager as the HILT.

3

Data Managers are positioned for Success

The DM skill set which includes, domain expertise, adaptability, communication, ethics, quality assurance, collaboration, feedback, documentation, and training are key to transitioning to Clinical Data Science.



Q & A

THANK YOU

Achilleas Zaras
Senior Manager of Solution Consulting
eClinical Solutions
azaras@eclinicalsol.com
eclinicalsol.com



PHUSE EU Connect 2023



Use Case #1: Anomalous AE Durations

Anomalous Subjects for Urinary tract infection 2

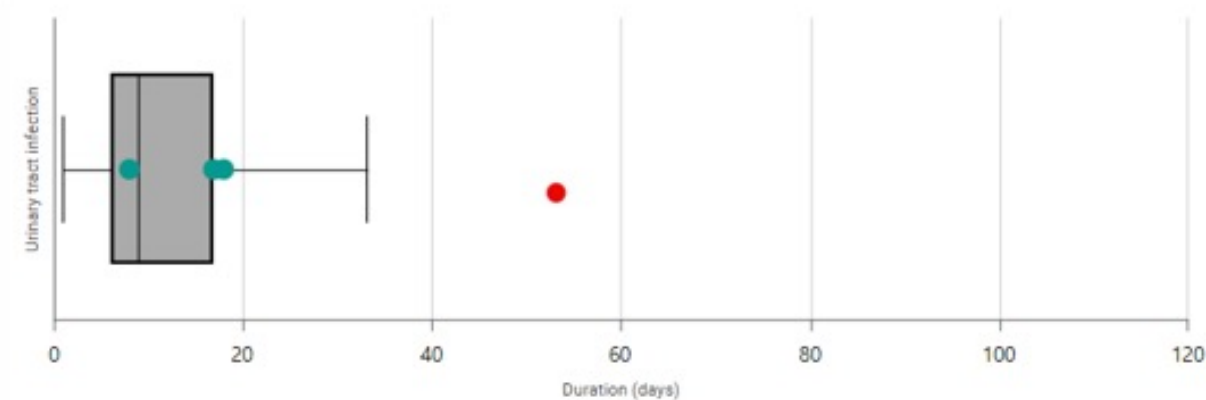
Review

☐	MDR	Subject ID	# of Events	# of Anomalous Events	Site ID
☒		CLINTEK-061-0789-808R	4	1	0789
☐		CLINTEK-061-8030-807R	1	1	8030

Medical History 29

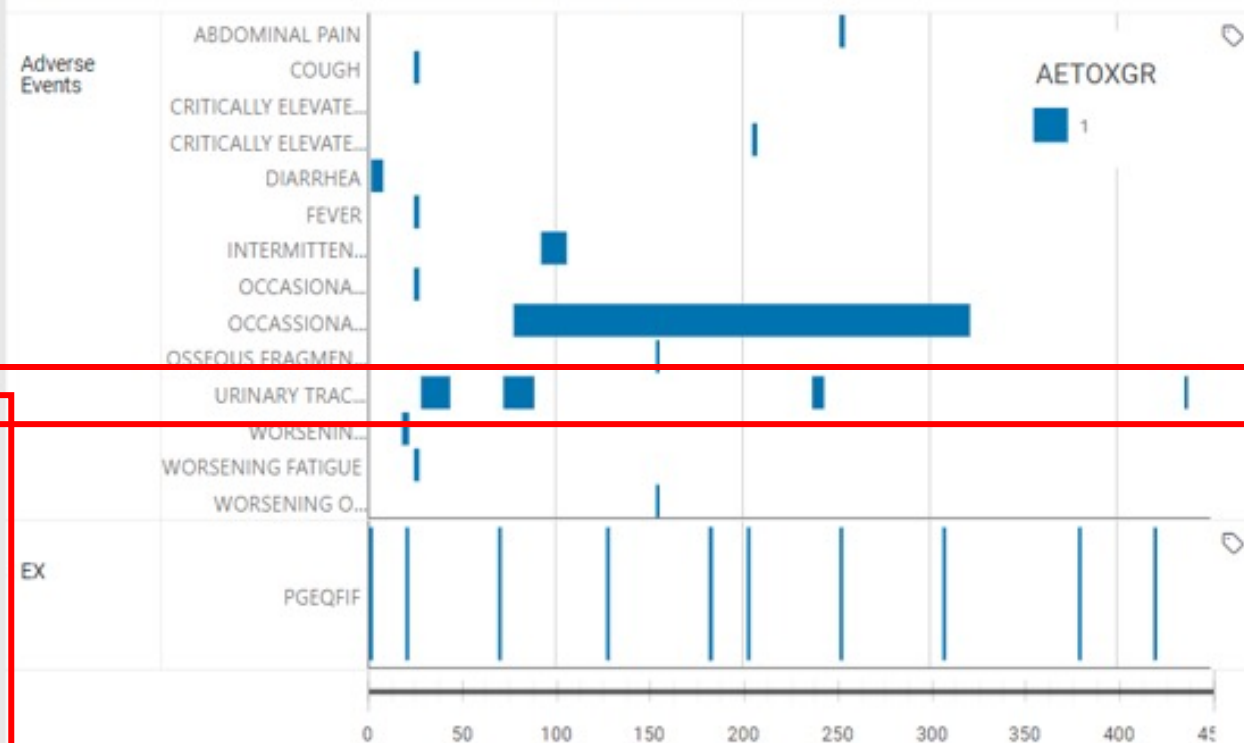
MHSEQ	MHTERM	MHDECOD	MHCAT	MHI
1	MYASTHENIA GRAVIS		MYASTHENIA GRAVIS MEDICAL ...	Y
2	RIGHT KNEE PAIN	Arthralgia	MEDICAL HISTORY	
3	LEFT BASILAR AIRSPACE DISEASE	Atelectasis	MEDICAL HISTORY	
4	ELEVATED GLUCOSE	Blood glucose increased	MEDICAL HISTORY	
5	BILATERAL CARPAL TUNNEL SY...	Carpal tunnel syndrome	MEDICAL HISTORY	

AE Durations



Graphical Patient Profile

USUBJID	CLINTEK-061-0789-808R
AGE	52
SEX	F
RACE	WHITE
ARM	COHORT 0



1

✓ Review ✓



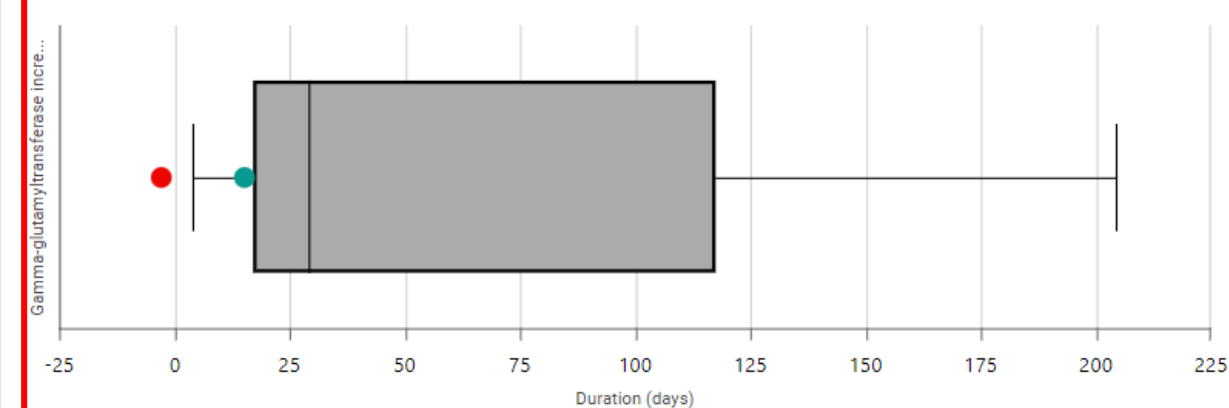
☒	DM	MDR	Subject ID	# of Events	# of Anomalous Events	Site ID
☒			CLINTEK-011-0407-008	2	1	080-0407

Medical History 4

4

MHSEQ	MHSPID	MHTERM	MHLLT	MHLLTCD	MHDECOD	MHPTCD	MHHLT
1	MH-926631-2	ALLERGIC RHINITIS	Allergic rhinitis	10001723	Rhinitis allergic	10039085	Nasal congestio
2	DH-926634-0	CLL DIAGNOSIS					
3	MH-926631-1	MENINGITIS	Meningitis	10027199	Meningitis	10027199	Central nervous
4	DH-926634-0	RT DIAGNOSIS					

AE Durations



Summary of trained AEs for Gamma-glutamyltransferase increased

Not Anomalous

Anomalous

Graphical Patient Profile



USUBJID	CLINTEK-011-0407-008
AGE	41
SEX	F
RACE	WHITE
ARM	COHORT 1

[illegible]

RANGE:

0

45

Use Case #2: CSM Labs and Vitals

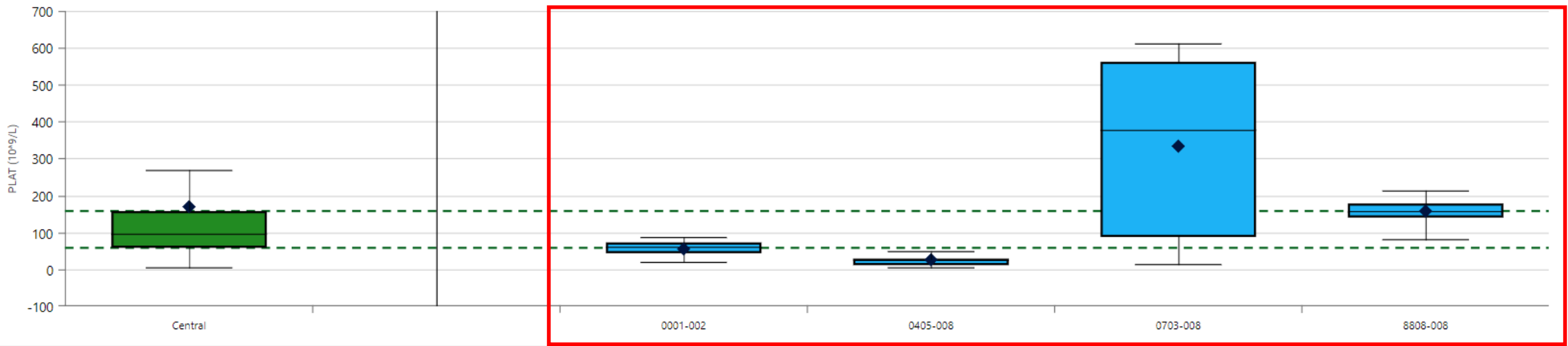
Lab and Vitals Tests Containing Anomalies 32

Domain	Test Category	Test Type	Units	# of Analyzed Subjects	# of Reviewed Subjects / # of Anomalous
LB	CHEMISTRY	GLUC	mmol/L	16	0 / 2
LB	HEMATOLOGY	HGB	g/L	16	0 / 8
VS	VS	HR	BEATS/MIN	15	0 / 4
LB	CHEMISTRY	K	mmol/L	15	0 / 2
LB	CHEMISTRY	LDH	ukat/L	16	0 / 7
LB	HEMATOLOGY	LYM	10^9/L	16	0 / 6
LB	HEMATOLOGY	MCH	pg/cell	16	0 / 7
LB	HEMATOLOGY	MCHC	g/L	16	0 / 5
LB	HEMATOLOGY	MCV	fL	16	0 / 6
LB	HEMATOLOGY	MONO	10^9/L	16	0 / 3
LB	HEMATOLOGY	NEUT	10^9/L	16	0 / 4
LB	HEMATOLOGY	NEUTB	10^9/L	6	0 / 3
LB	HEMATOLOGY	PLAT	10^9/L	16	0 / 4

Anomalous Subjects for PLAT 4

☐	DM	MDR	Subject ID	# of Analyzed Measurements	Site ID
☐			CLINTEK-011-0001-002	8	080-0001
☐			CLINTEK-011-0405-008	5	080-0405
☐			CLINTEK-011-0703-008	6	080-0703
☐			CLINTEK-011-8808-008	13	080-8808

Central Subject vs Anomalous Subjects



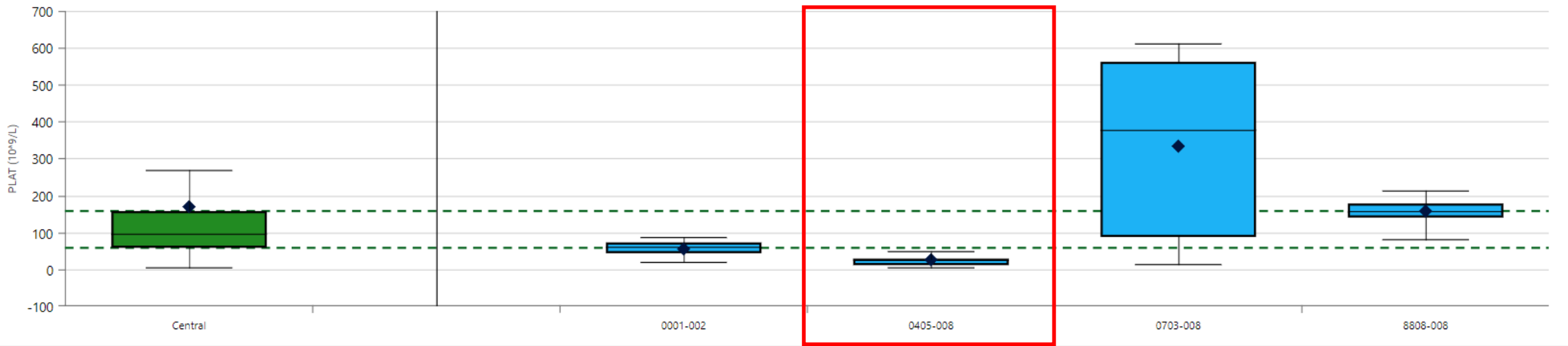
Lab and Vitals Tests Containing Anomalies 32

Domain	Test Category	Test Type	Units	# of Analyzed Subjects	# of Reviewed Subjects / # of Anomalous
LB	CHEMISTRY	GLUC	mmol/L	16	0 / 2
LB	HEMATOLOGY	HGB	g/L	16	0 / 8
VS	VS	HR	BEATS/MIN	15	0 / 4
LB	CHEMISTRY	K	mmol/L	15	0 / 2
LB	CHEMISTRY	LDH	ukat/L	16	0 / 7
LB	HEMATOLOGY	LYM	10^9/L	16	0 / 6
LB	HEMATOLOGY	MCH	pg/cell	16	0 / 7
LB	HEMATOLOGY	MCHC	g/L	16	0 / 5
LB	HEMATOLOGY	MCV	fL	16	0 / 6
LB	HEMATOLOGY	MONO	10^9/L	16	0 / 3
LB	HEMATOLOGY	NEUT	10^9/L	16	0 / 4
LB	HEMATOLOGY	NEUTB	10^9/L	6	0 / 3
LB	HEMATOLOGY	PLAT	10^9/L	16	0 / 4

Anomalous Subjects for PLAT 4

☐	DM	MDR	Subject ID	# of Analyzed Measurements	Site ID
☐			CLINTEK-011-0001-002	8	080-0001
☐			CLINTEK-011-0405-008	5	080-0405
☐			CLINTEK-011-0703-008	6	080-0703
☐			CLINTEK-011-8808-008	13	080-8808

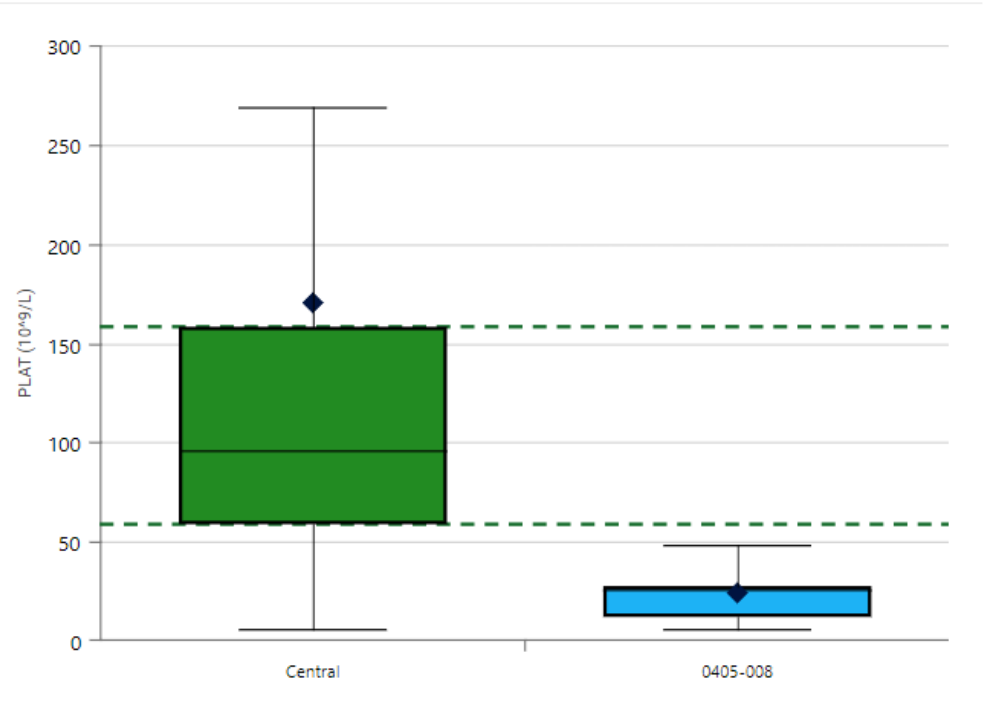
Central Subject vs Anomalous Subjects



Anomalous Subjects for PLAT 4

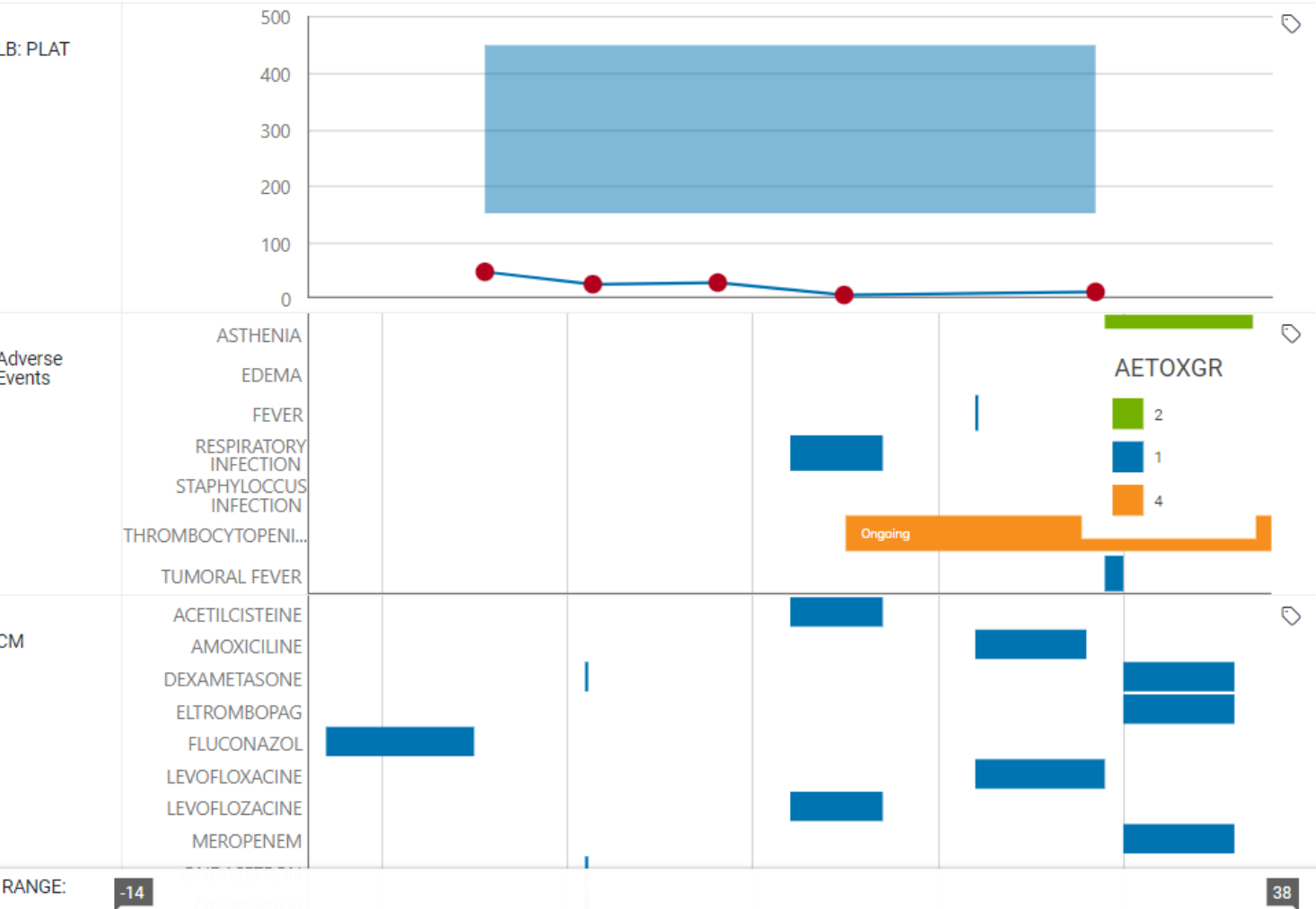
☒	DM	MDR	Subject ID	# of Analyzed Measurements	Site ID
☐			CLINTEK-011-0001-002	8	080-0001
☒			CLINTEK-011-0405-008	5	080-0405
☐			CLINTEK-011-0703-008	6	080-0703
☐			CLINTEK-011-8808-008	13	080-8808

Central Subject vs Anomalous Subjects



Graphical Patient Profile

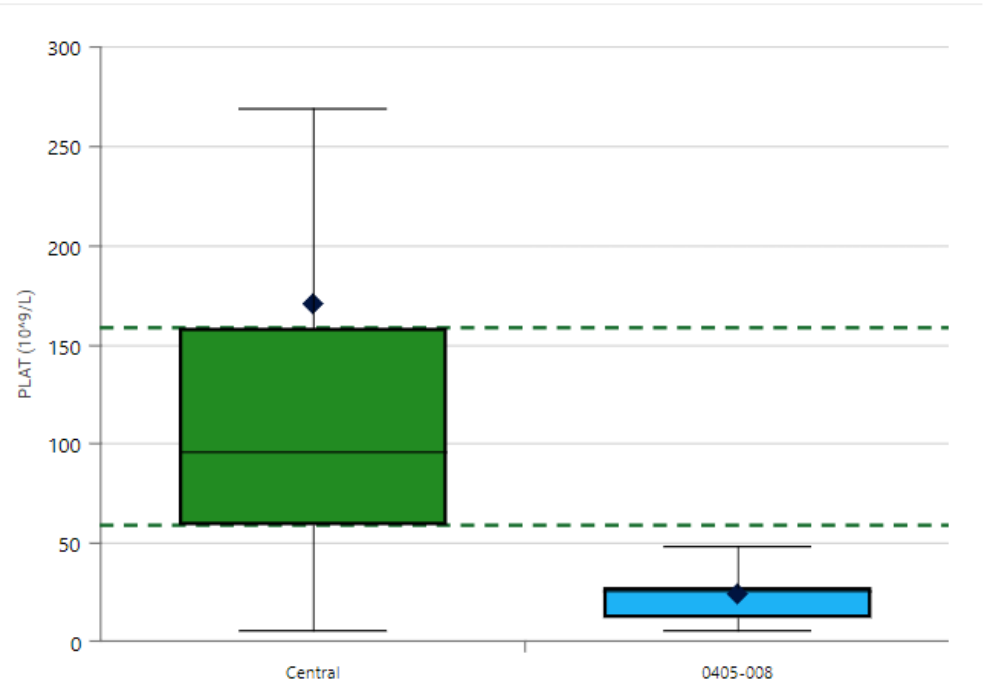
USUBJID	CLINTEK-011-0405-008
AGE	68
SEX	M
RACE	WHITE
ARM	COHORT 1



Anomalous Subjects for PLAT 4

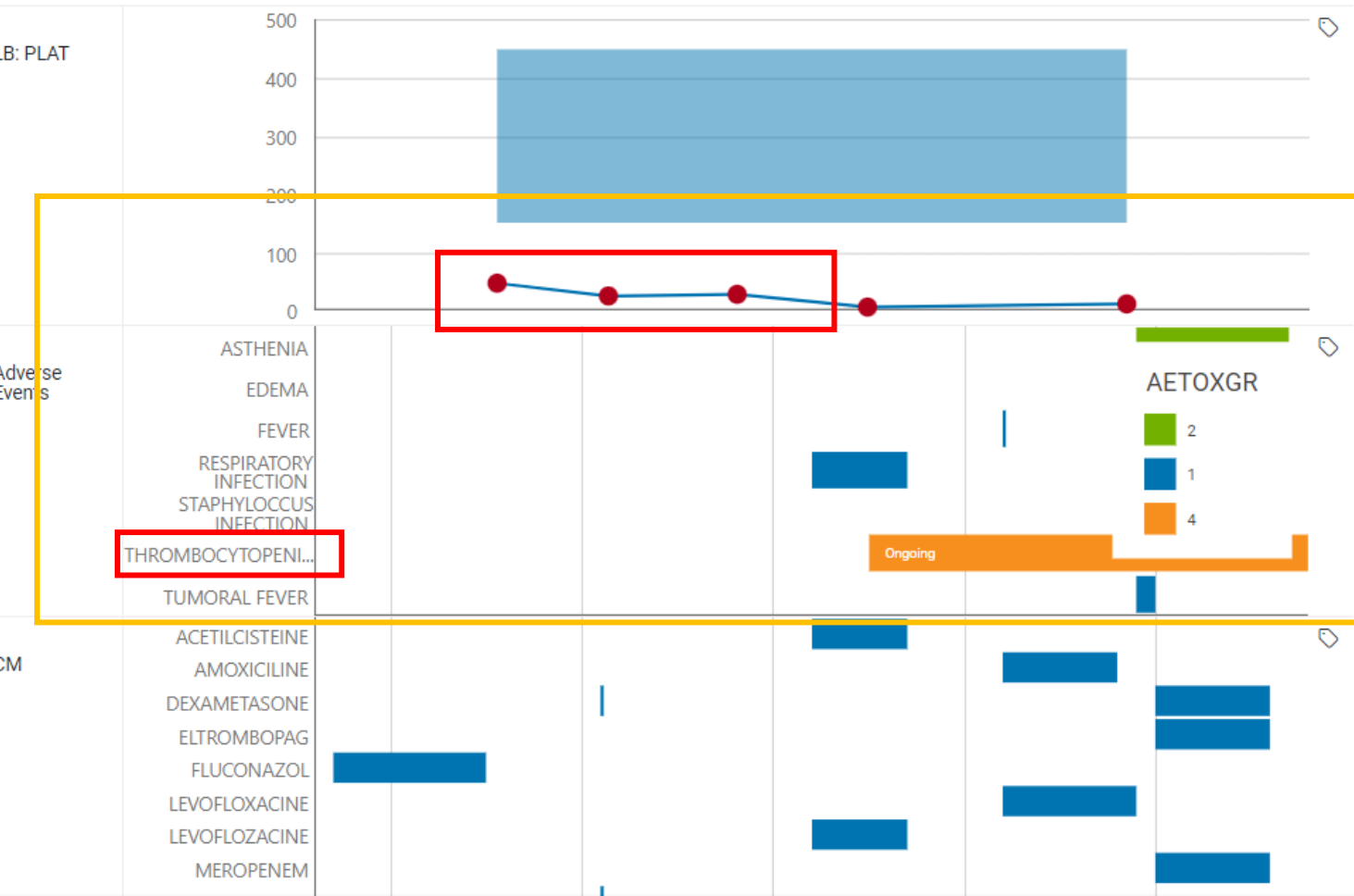
☐	DM	MDR	Subject ID	# of Analyzed Measurements	Site ID
☐			CLINTEK-011-0001-002	8	080-0001
☒			CLINTEK-011-0405-008	5	080-0405
☐			CLINTEK-011-0703-008	6	080-0703
☐			CLINTEK-011-8808-008	13	080-8808

Central Subject vs Anomalous Subjects



Graphical Patient Profile

USUBJID	CLINTEK-011-0405-008
AGE	68
SEX	M
RACE	WHITE
ARM	COHORT 1



Use Case #3: Domain Classification

Subjects

276

276

Issues

1

6

eIQ Review

CSM Labs and Vitals

AE Durations

Shift in Labs and Vitals

Univariate Outlier Detection

Incorrect Item Type

CM Durations

Domain Classification

CM Consistency

Visualizations

+

>

Data

>

Domain Classification > Clinical-AE > Record: 1033

Tests Containing Anomalies 3

Data Store	Domain	# of Analyzed Subjects	# of Analyzed Records	# of Anomalous Subjects
Reporting	MH	202	3619	122
Clinical	AE	208	4605	1
Clinical	MH	202	3418	124

Anomalous Records for Clinical-AE 1

+

Review

☒	DM	MM	Record ID	Subject ID	Predicted Domain	Misclassified Variables	Reason
☒			1033	7010	Procedure	1	The re

AE for Subject 7010 14

	DM	MM			PROJECT	ENVIRONMENTNAME	SUBJECT	SITE
0	0	0	0	0	CLINTEK-012	Prod	0008-007	0008 - H.
0	0	0	0	0	CLINTEK-012	Prod	0008-007	0008 - H.
0	0	0	0	0	CLINTEK-012	Prod	0008-007	0008 - H.
0	0	0	0	0	CLINTEK-012	Prod	0008-007	0008 - H.
0	0	0	0	0	CLINTEK-012	Prod	0008-007	0008 - H.
0	0	0	0	0	CLINTEK-012	Prod	0008-007	0008 - H.

Record Details

RECORDPOSITION	12
MINCREATED	42524.8531597222
MAXUPDATED	43688.7245717593
SAVETS	43716.9148032407
CODER_HIERARCHY	
RECORDACTIVE	
AETERM	RVP HUMAN RHINOVIRUS
AETERM_HGT	Microbiology and serology investigations
AETERM_HGTN	10027529
AETERM_HLT	Virus identification and serology
AETERM_HLTN	10047490
AETERM_LLT	Human rhinovirus test
AETERM_LLTN	10075163
AETERM_PT	Human rhinovirus test

Based on Historic data, AETERM should very likely be recorded under Procedure Domain.

38

Subjects

276

276

Issues

1

6

eIQ Review

CSM Labs and Vitals

AE Durations

Shift in Labs and Vitals

Univariate Outlier Detection

Incorrect Item Type

CM Durations

Domain Classification

CM Consistency

Visualizations

+

>

Data

>

Domain Classification > Clinical-MH > Record: 11900

Tests Containing Anomalies

3

Data Store	Domain	# of Analyzed Subjects	# of Analyzed Records	# of Anomalous Subjects
Reporting	MH	202	3619	122
Clinical	AE	208	4605	1
Clinical	MH	202	3418	124

Anomalous Records for Clinical-MH

346

+

Review

☐	DM	MM	Record ID	Subject ID	Predicted Domain	Misclassified Variables	Re
☒			11900	10016	Procedure	1	The
☐			6646	10163	Procedure	1	The
☐			6397	10228	Procedure	1	The
☐			7989	10228	Procedure	1	The
☐			9140	10373	Procedure	1	The
☐			8990	10378	Procedure	1	The
☐			7602	10378	Procedure	1	The
☐			8398	10378	Procedure	1	The
☐			6480	10378	Procedure	1	The
☐			6842	10378	Procedure	1	The
☐			6906	10378	Procedure	1	The
☐			12279	10378	Procedure	1	The
☐							

MH for Subject 10016

20

	DM	MM			PROJECT	ENVIRONMENTNAME	SUBJECT	SITE
0	0	0	0	0	CLINTEK-012	Prod	0071-089	0071 - AI
0	0	0	0	0	CLINTEK-012	Prod	0071-089	0071 - AI
0	0	0	0	0	CLINTEK-012	Prod	0071-089	0071 - AI
0	0	0	0	0	CLINTEK-012	Prod	0071-089	0071 - AI
0	0	0	0	0	CLINTEK-012	Prod	0071-089	0071 - AI
0	0	0	0	0	CLINTEK-012	Prod	0071-089	0071 - AI
0	0	0	0	0	CLINTEK-012	Prod	0071-089	0071 - AI

Record Details

+

RECORDPOSITION	8
MINCREATED	42779.6903009259
MAXUPDATED	43684.0289351852
SAVETS	43716.9242476852
CODER_HIERARCHY	
RECORDACTIVE	1
MHYN	
MHYN_STD	
MHTERM	KNEE ARTHROSCOPY
MHTERM_HGT	Musculoskeletal and soft tissue investigations (e enzyme tests)
MHTERM_HGTN	10028387
MHTERM_HLT	Musculoskeletal and soft tissue imaging procedu
MHTERM_HLTN	10028386
MHTERM_LIT	Knee arthroscopy

Based on Historic data, MHTERM should very likely be recorded under Procedure Domain.

39