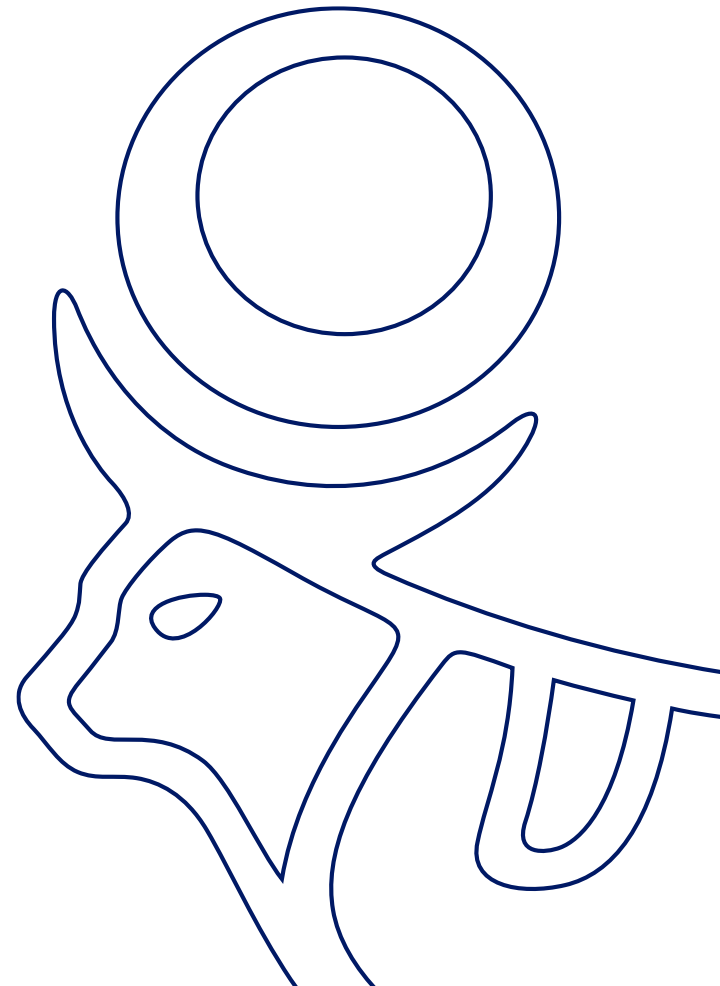


Potency Assay Generated Data - Is It a Minefield or a Knowledge Bank and Are They Useable

What's our approach to dealing with data?

Jan Amstrup, Ph.D.

29 April 2026



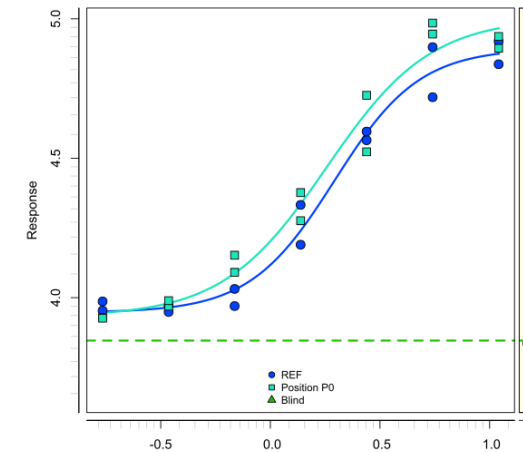
Visualize data in real time

Machine Learning to predict outcome of an assay

Aggenda

Speed is the key to success – or is it?

- Frequently, assays are portrayed as successful and the problems are rarely disclosed



- We can't keep simply hiring people
So, we need to make life easier for ourselves



Speed is the Key to Success – or is it?

- Potency assays produce large, complex datasets, but laboratories often lack the time to analyse, trend, and interpret this data
- Data analysis — not data generation — has become the main bottleneck in laboratories and day-to-day operations
- Daily operations have the highest priority; yet re-runs are often described as a waste of time—and therefore waste of money
- Data is created and stored using different IT systems 🙄
- One of the hottest things you can become these days is a Data Scientist – or AI JeDi master

How can Data be used?

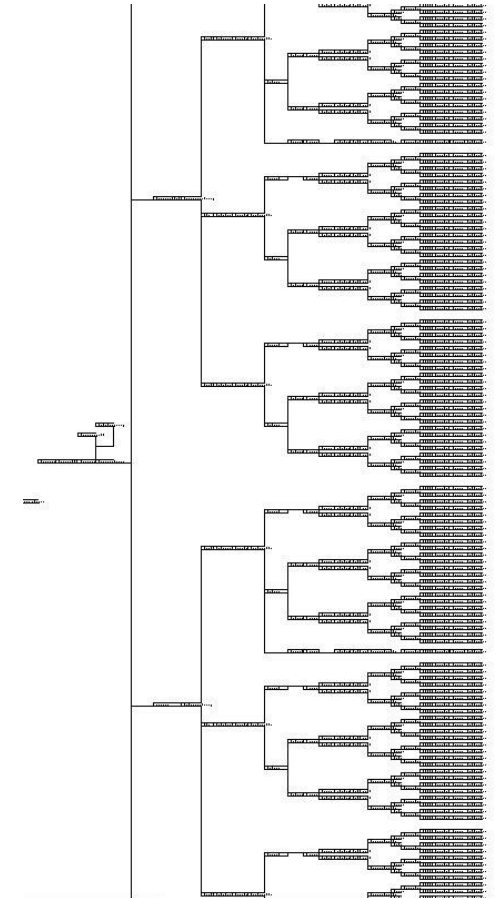
- Two approaches
- Visualize data in real time
 - Various pitfalls — what to watch for and how to interpret the data
- Machine Learning to predict outcome of an assay
 - What volume of data is needed to achieve precise predictions, and what level of precision is necessary?

Visualisation of Data

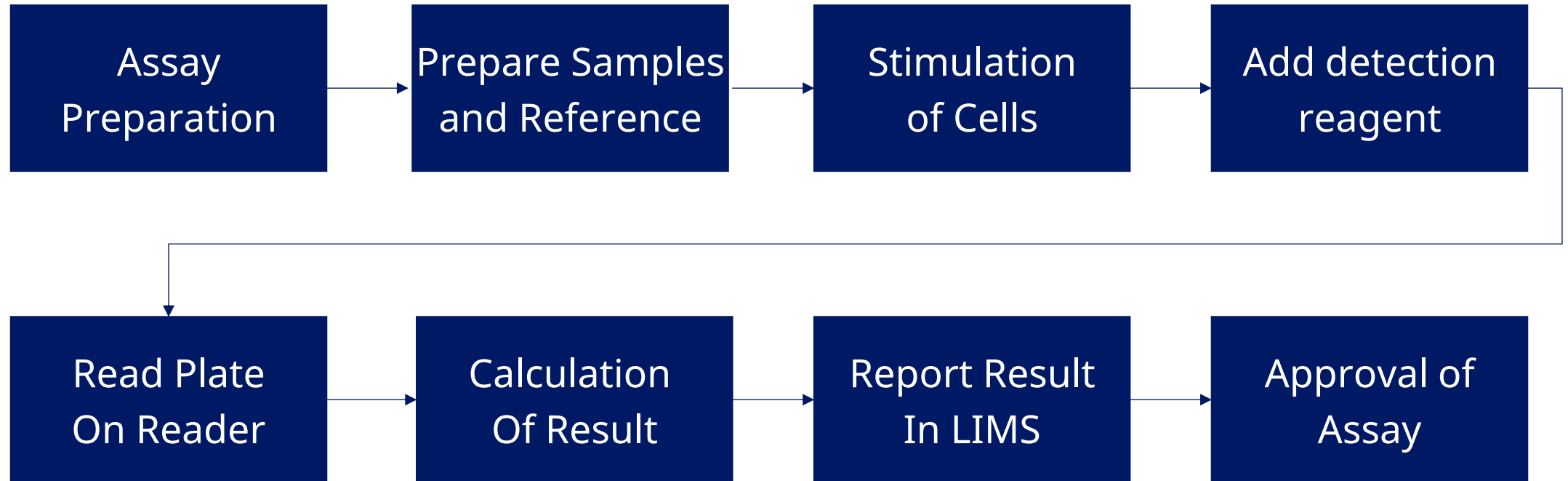
- Is the assay performing as wanted?
- Where is the data stored?
- How do we find the pitfalls of the assay and the data?

Complex Process in Laboratory - How it Started

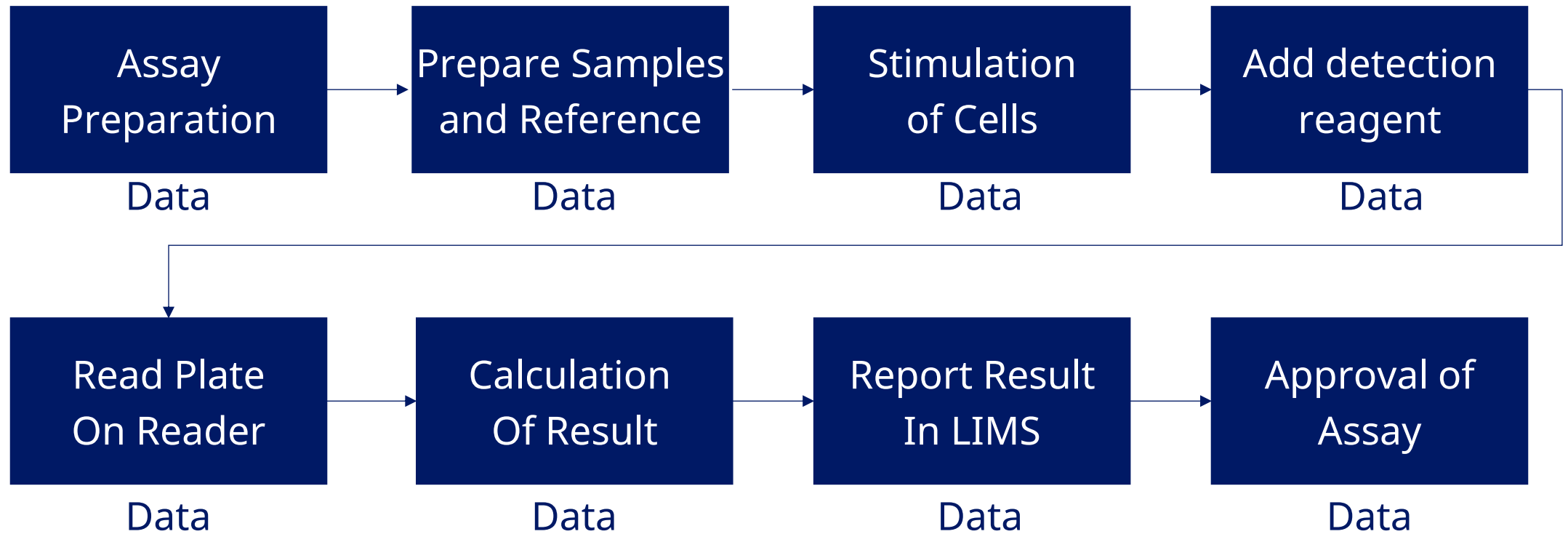
- A SOP should not be so strict, needs **Freedom to Operate**. Otherwise, deviations are generated
- Lab technicians do what they usually do and what they find appropriate
- Process mapping of an assay to standardize without loosing Freedom of Operation



Assay Overview – Cell Based Assay



Assay Overview – Cell Based Assay



How we look at data - We use Power BI in Real Time

- Integrates data from diverse sources
- Monitor assay performance from different angles
- Easily identify outliers and spot alarming trends
 - Recognize negative trends as well as outliers among the categories
- Explore details and find problem root causes
- Dashboards can be tailored to different user roles

- Data is collected from LIMS

Content Dashboard



Content Sample ID	Content Sample Status	Content	Content Method	Bioactivity Sample ID	Bioactivity Sample Status	Bioactivity	Bioactivity Method
S-20260413-00129	InProgress	58,62660222219	0708930_mg/ml_korr	S-20260413-00128	InProgress	63,37155788891	
S-20260413-00129	InProgress	58,62660222219	0708930_mg/ml_korr	S-20260413-00128	InProgress	28,34271	
S-20260328-00424	InProgress		0812130_mg/mg	S-20260328-00421	InProgress	,8459135914736	
S-20260328-00662	InProgress		0812130_mg/mg	S-20260328-00658	InProgress	,8881610881296	
S-20260331-00173	InProgress		0812130_mg/mg	S-20260331-00170	InProgress	,8809624459557	
S-20260331-00186	InProgress		0812130_mg/mg	S-20260331-00185	InProgress	,9112019413726	
S-20260328-00424	InProgress		0812130_mg/mg	S-20260328-00421	InProgress	,8880521506832	
S-20260328-00662	InProgress		0812130_mg/mg	S-20260328-00658	InProgress	,8756758593446	
S-20260331-00173	InProgress		0812130_mg/mg	S-20260331-00170	InProgress	,8344401456962	
S-20260331-00186	InProgress		0812130_mg/mg	S-20260331-00185	InProgress	,8532220366709	
S-20260325-00144	Reviewed	0,4723326557241	0777124_DPiv_mg/ml	S-20260325-00141	InProgress	,48521	
S-20260413-00108	InProgress	185,33252633	0708930_mg/ml_wt	S-20260413-00107	Received		
S-20260411-00027	InProgress	0,8036890088	0980800_MG/MG	S-20260411-00026	InProgress	,77272	

Different assays

Bioactivity Method
All

Platform
All

Date
1/1/2026 4/16/2026

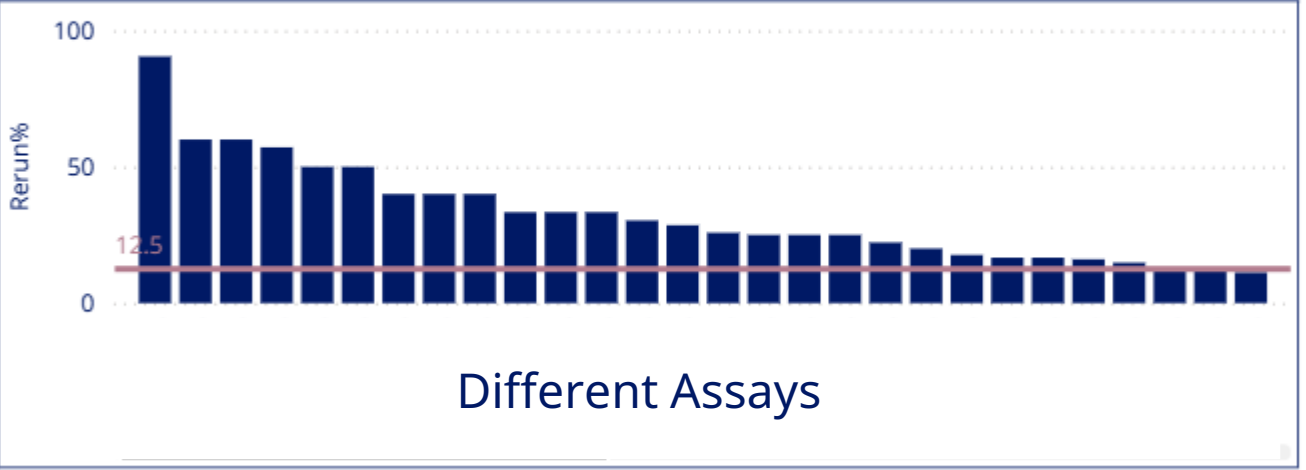
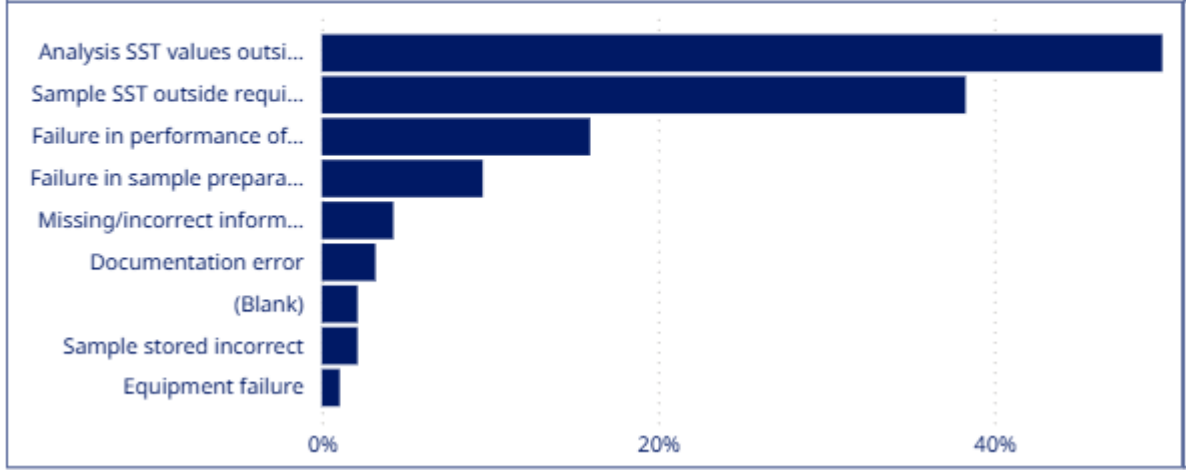
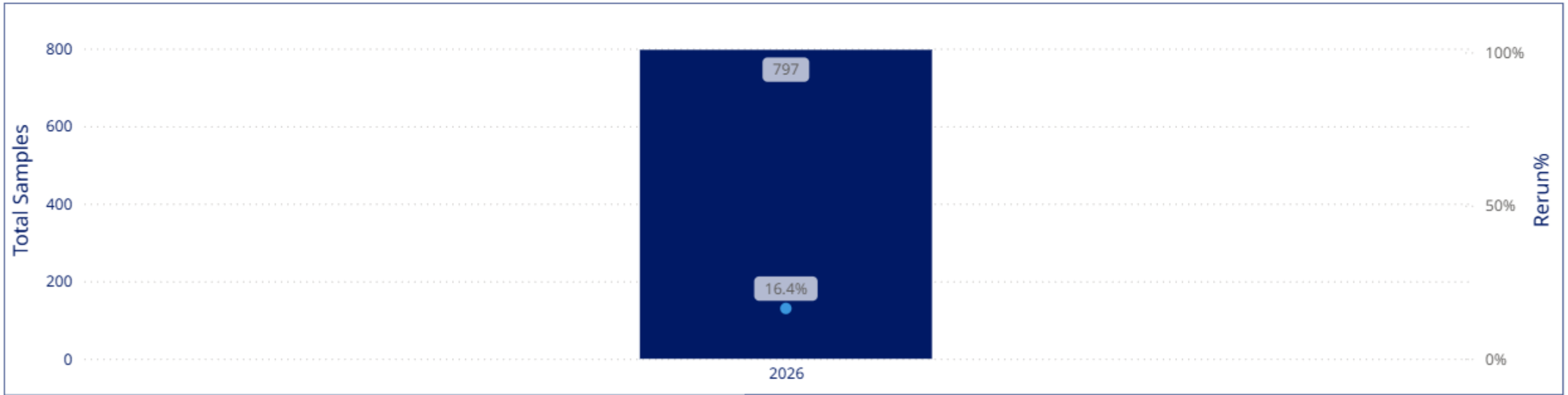
Year
Month
Week

Rerun%
16.44 %

Rerun Samples
131

Unique Samples
688

Total Samples
797



Bioactivity Method

All

Platform

All

Date

1/1/2026

4/16/2026

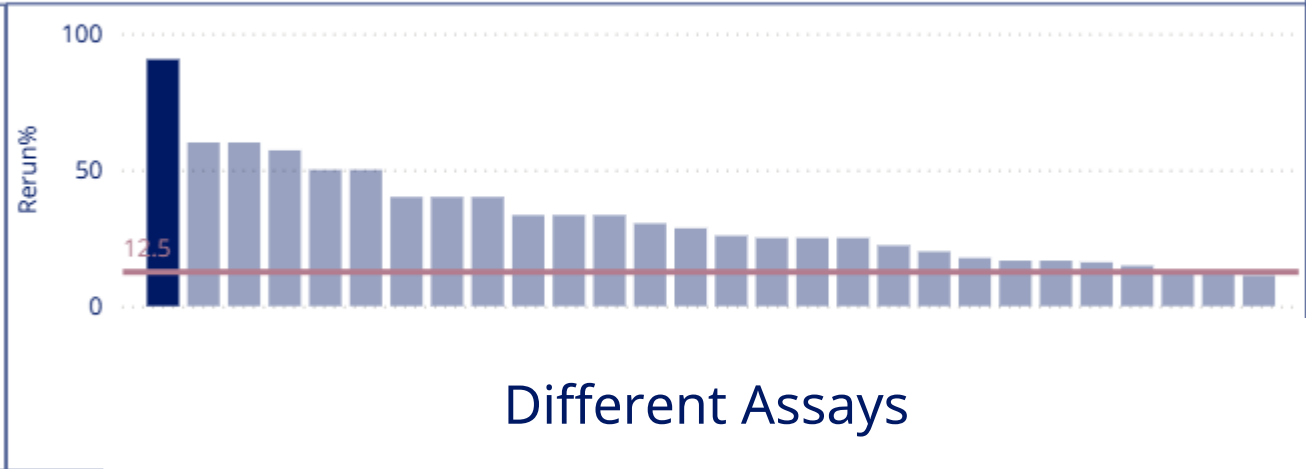
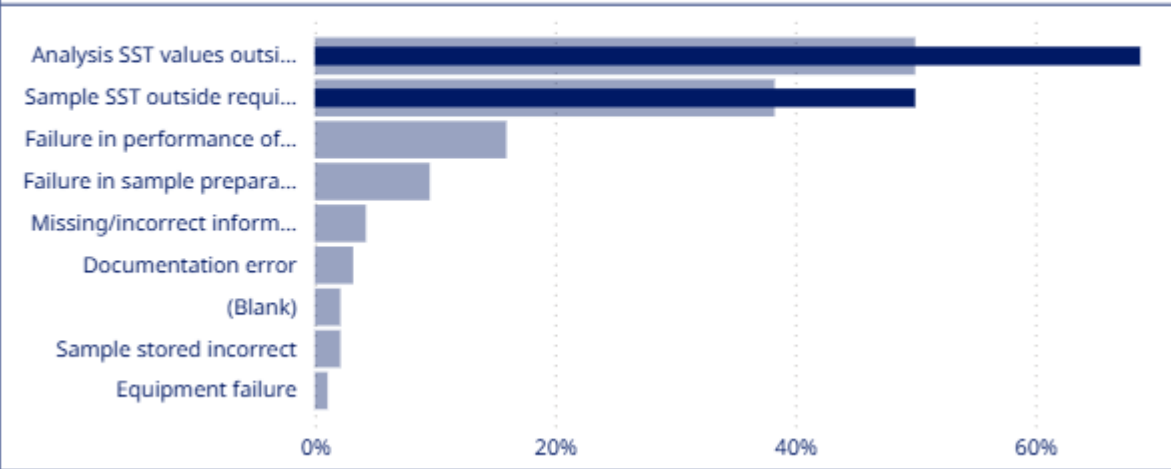
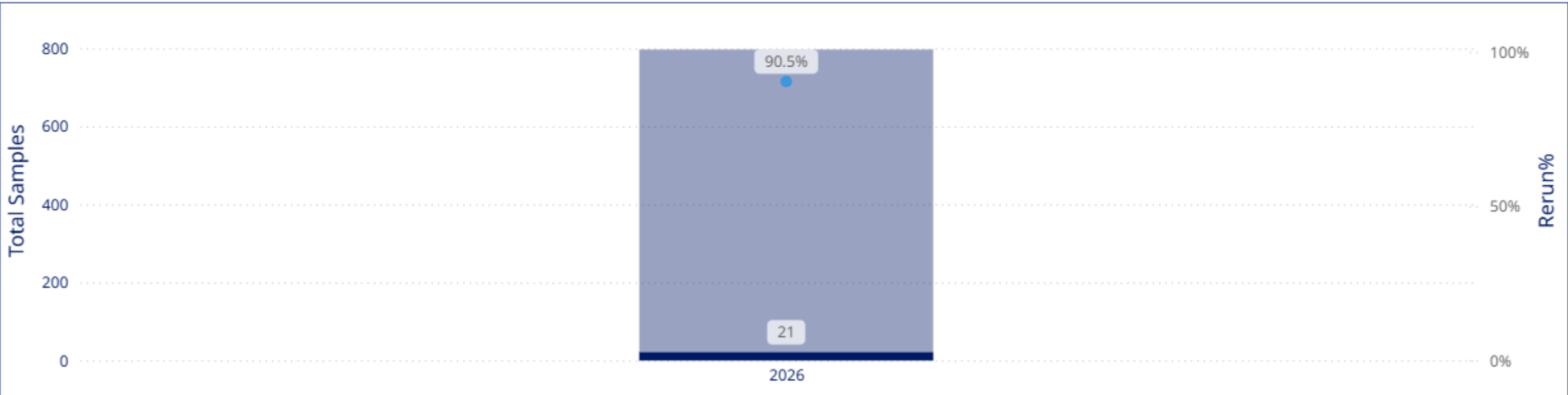
- Year
- Month
- Week

Rerun%
90.48 %

Rerun Samples
19

Unique Samples
18

Total Samples
21



Different Assays

Bioactivity Method
 All

Platform
 All

Date
 1/1/2026 4/16/2026

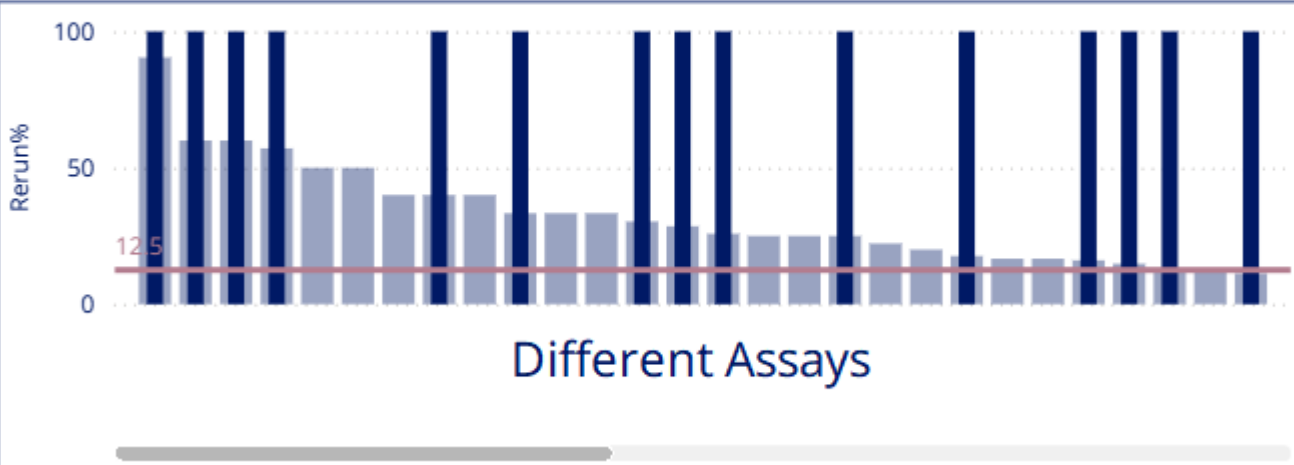
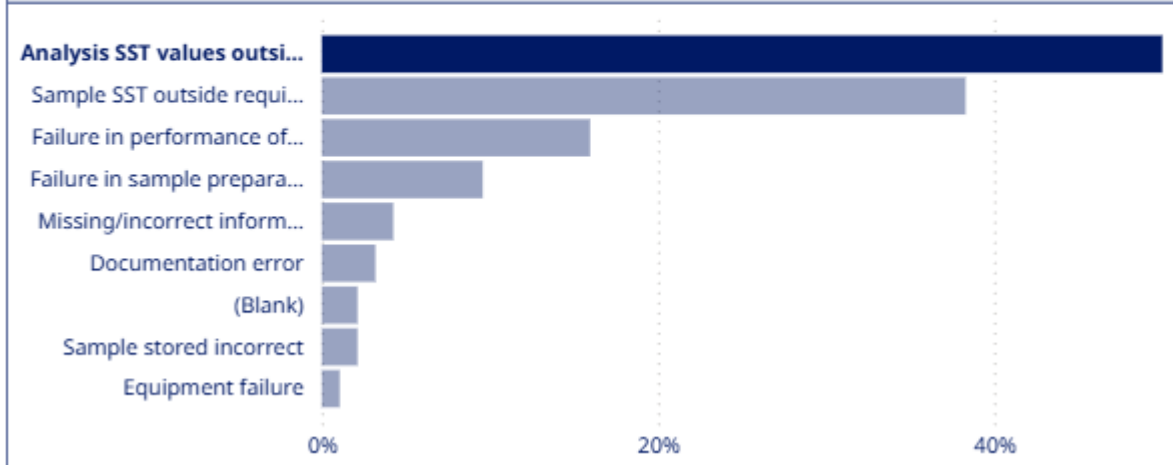
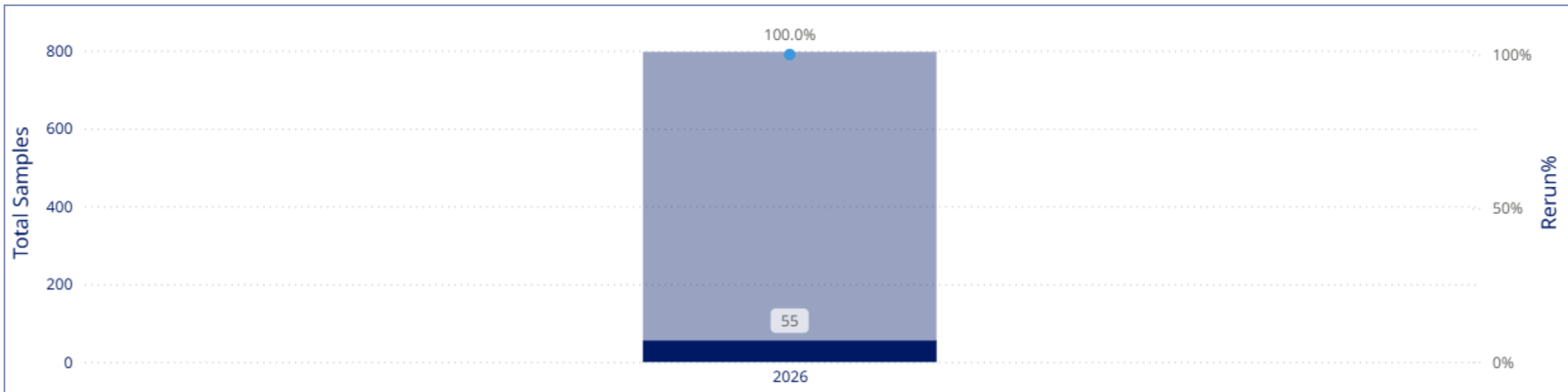
Year
 Month
 Week

Rerun%
 100.00 %

Rerun Samples
 55

Unique Samples
 48

Total Samples
 55



Assay Overview

NOVO NORDISK

Date

2025

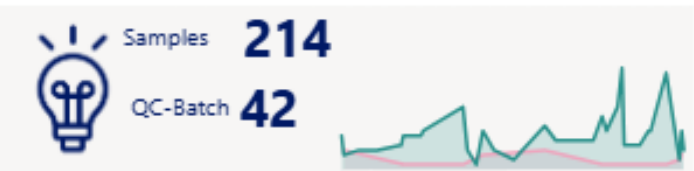
Sample Method

All

Biograph Project

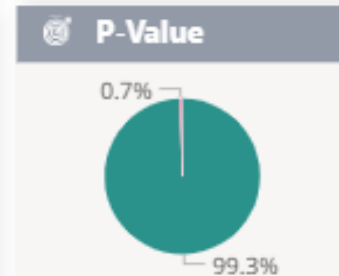
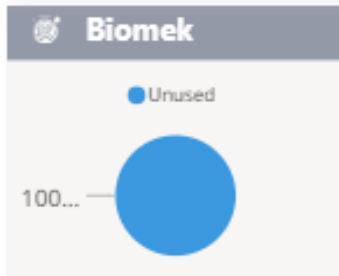
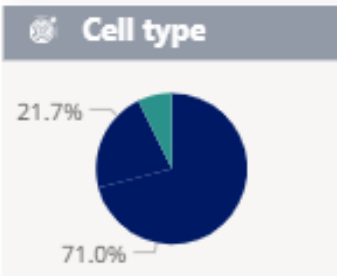
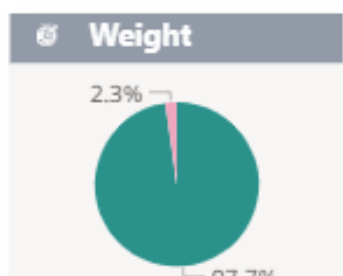
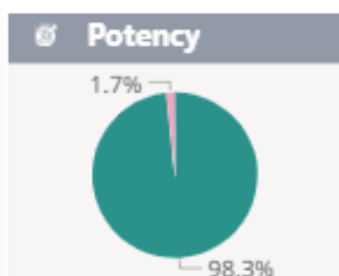
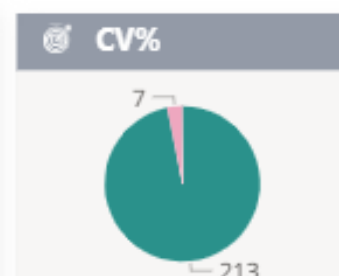
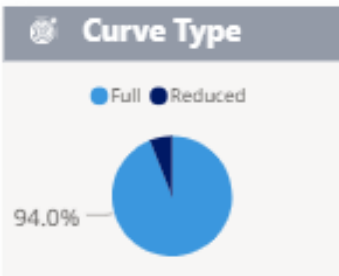
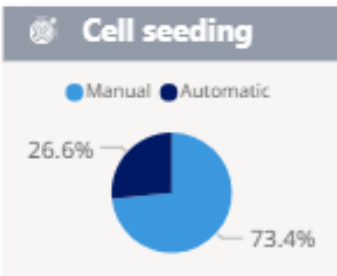
Total Sample	Rerun Samples	Rerun-%
214	21	9.8

Historical Data



Passed (All SST)	# Positions	% Position
False	31	4.43%
True	668	95.57%
Total	699	100.00%

Passed (All)	# Sample ID	% Sample ID
True	208	97.20%
False	25	11.68%
Total	214	100.00%



Manual Cell Seeding

Date: 2025 | Sample Method: All | Biograph Project: []

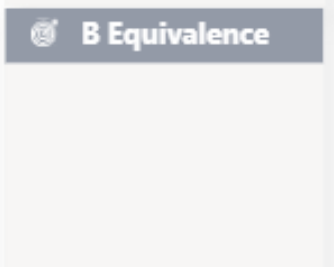
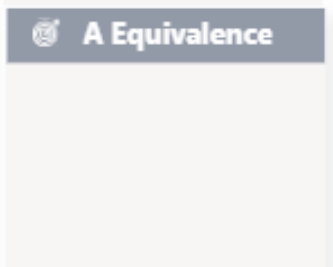
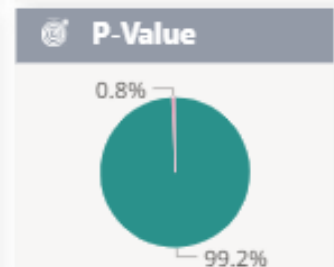
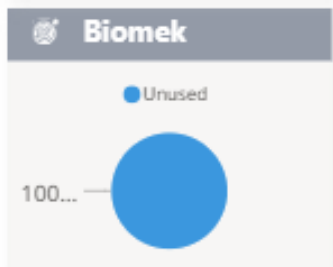
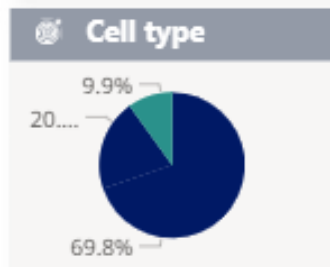
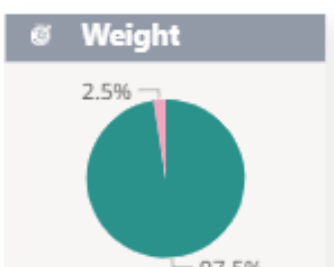
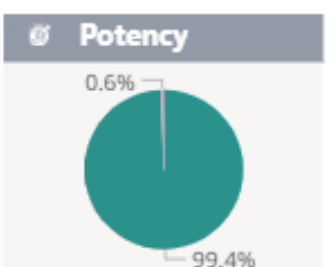
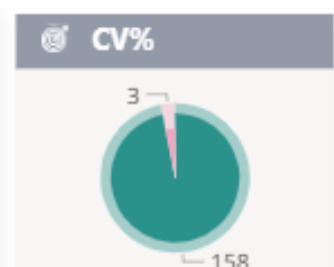
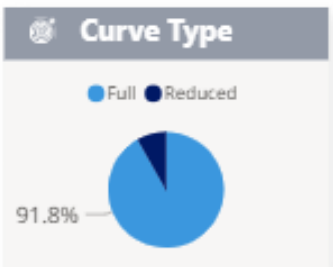
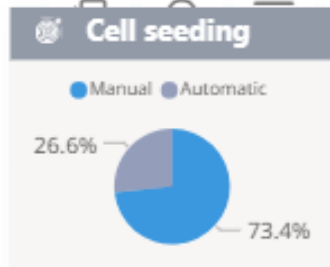
Total Sample	Rerun Samples	Rerun-%
160	17	10.6

Historical Data

Samples: **160**
QC-Batch: **32**

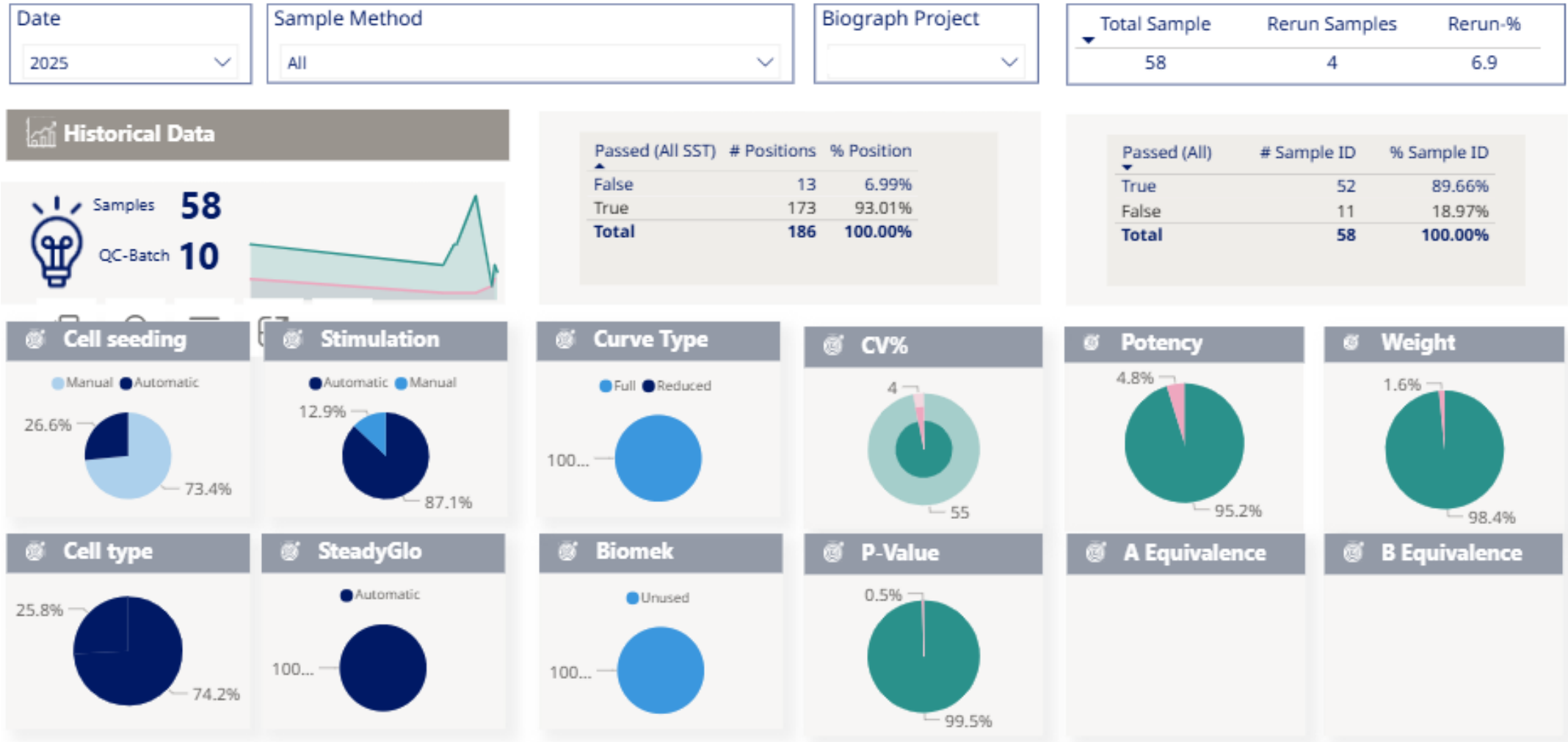
Passed (All SST)	# Positions	% Position
False	18	3.51%
True	495	96.49%
Total	513	100.00%

Passed (All)	# Sample ID	% Sample ID
True	158	98.75%
False	14	8.75%
Total	160	100.00%

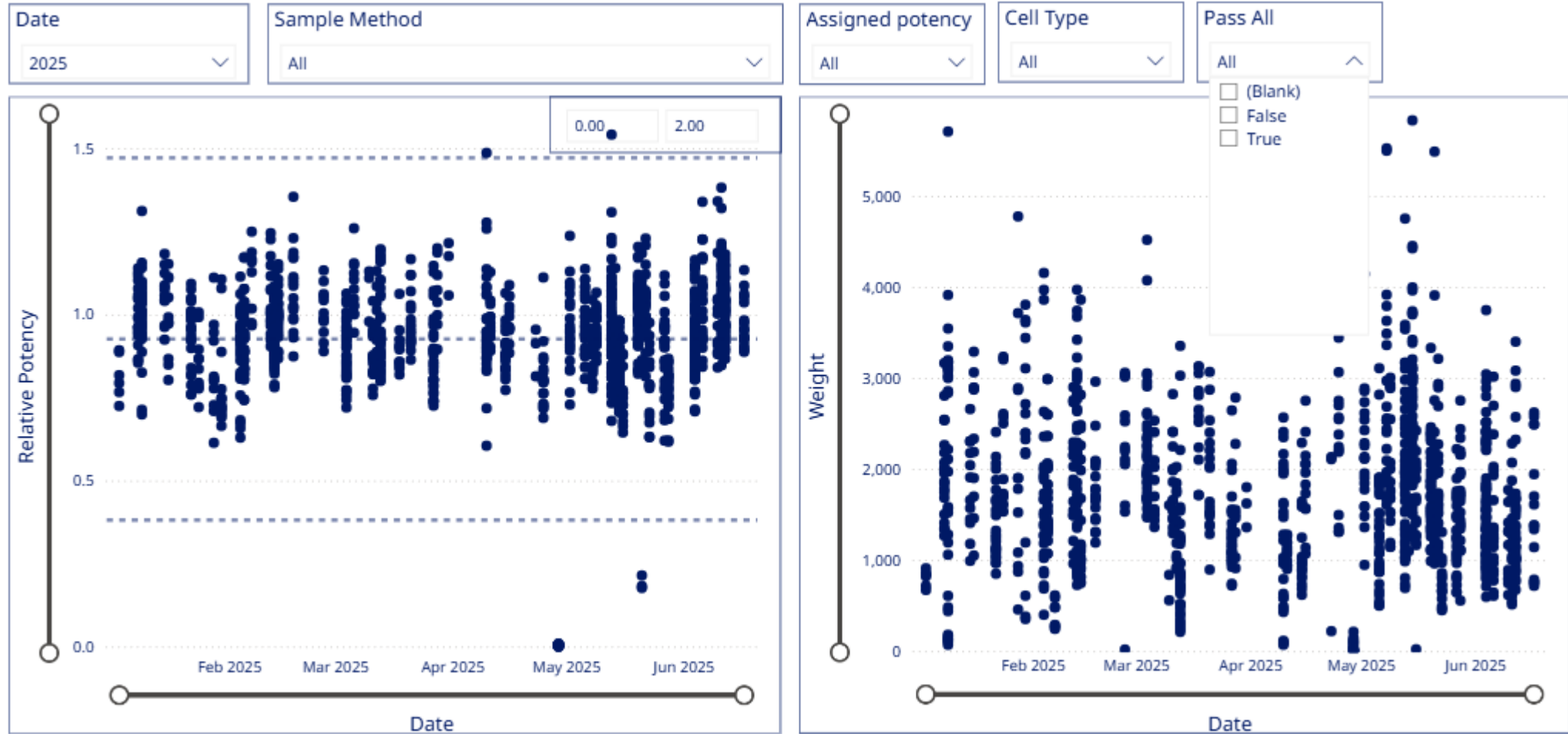


Automatic Cell Seeding

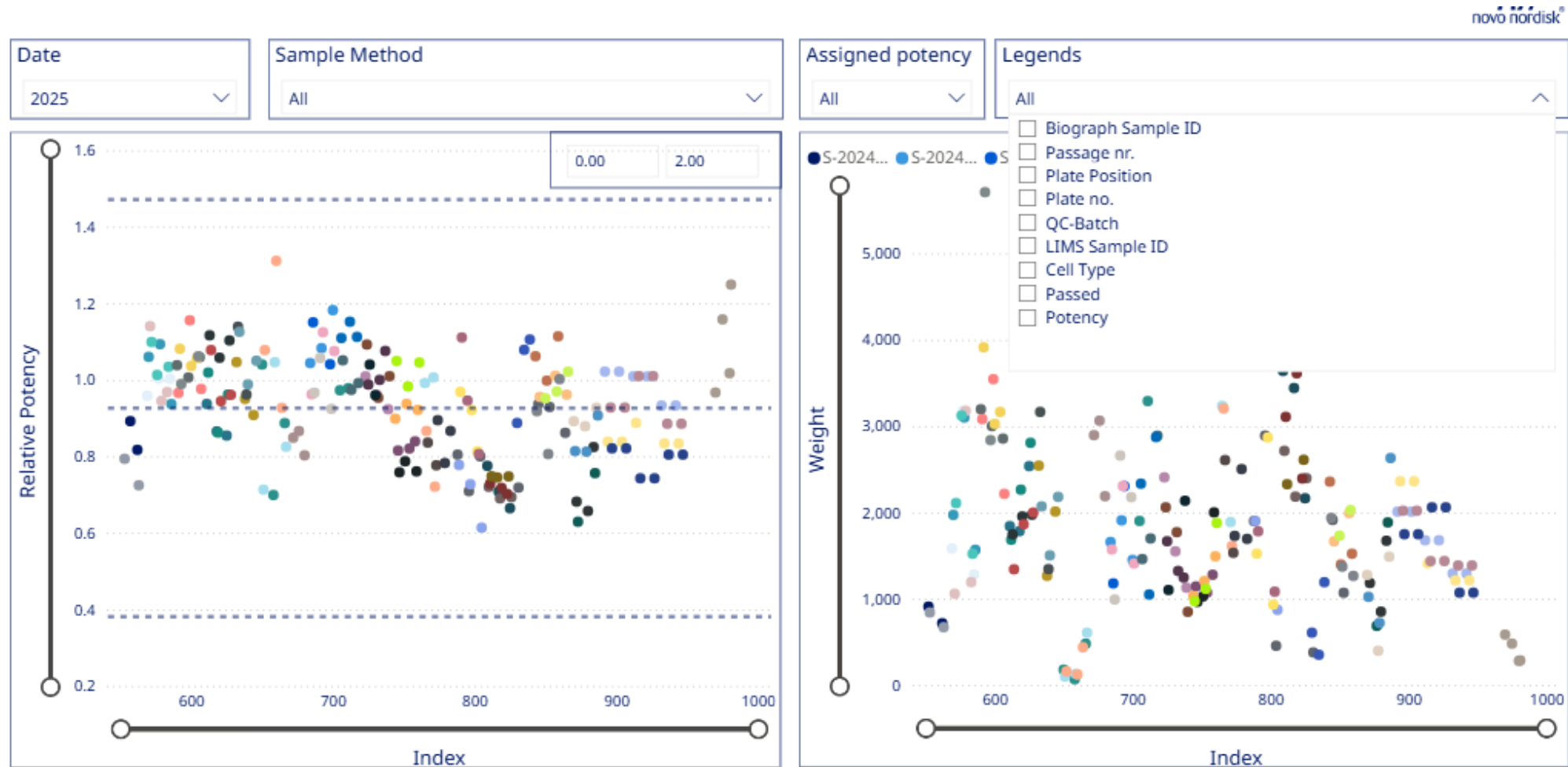
novo nordisk



Assay Trending - Date



Assay Trending - Index



Search

Filters on this visual

Biograph Sample ID ^

is (All)

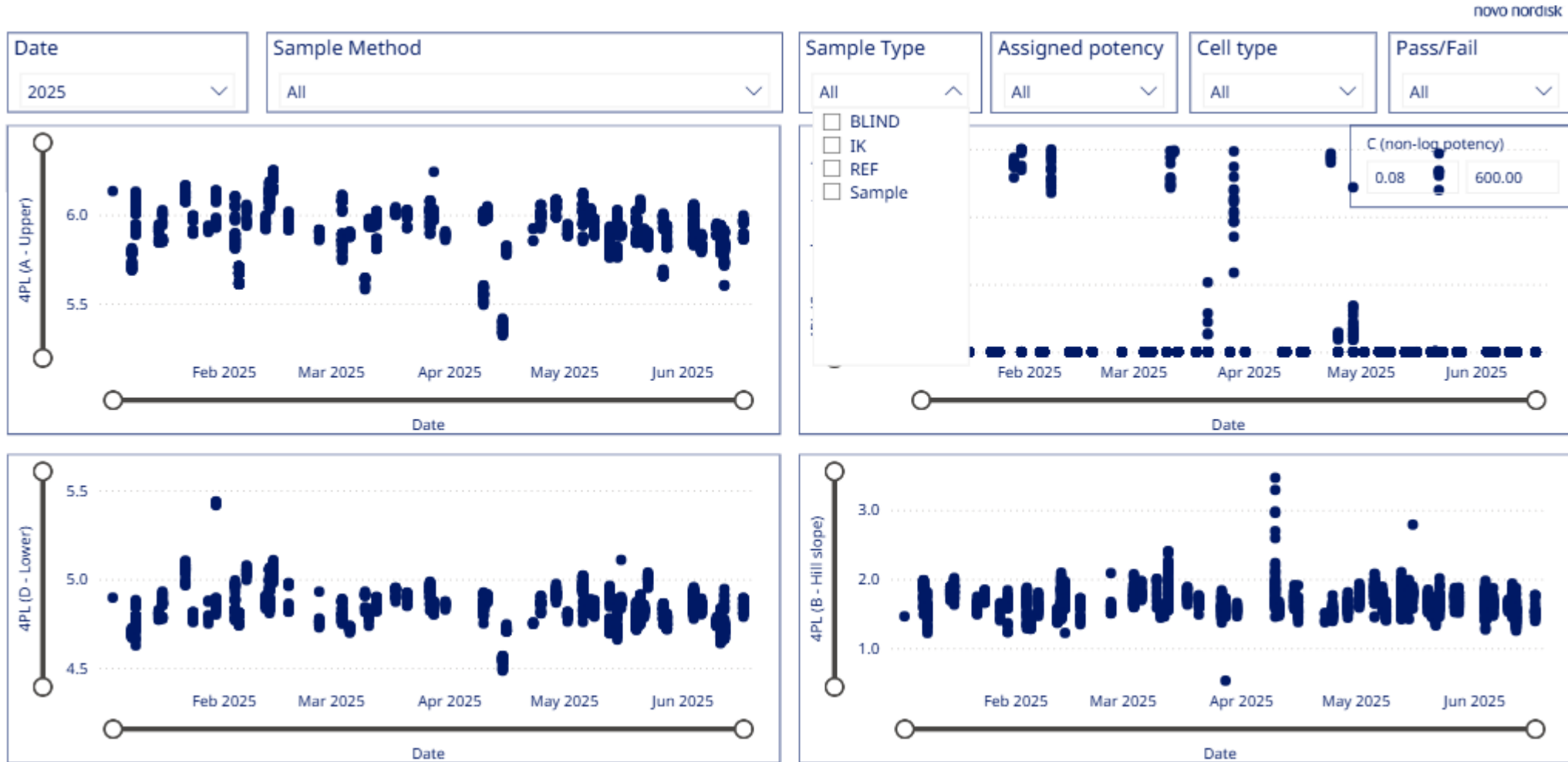
Filter type

Basic filtering

Search

- BLIND-QB-20250611... 3
- BLIND-QB-20250612... 7
- BLIND-QB-20250613... 4
- IK-QB-20241230-00... 3
- IK-QB-20250107-00... 6
- IK-QB-20250108-00... 6

4PL Trending - Date



Pitfalls with the Data

- Reruns are not recorded consistently among lab technicians and chemists.
 - Definition of reruns. Need for standardisation
- Registration of how data are generated
 - Which robots are use?
- Fail on SST parameters, and which?
- Are small deviation throughout the assay documented? And how?
 - Important to standardise the assay and the data?

So, What Else can Data be used for?

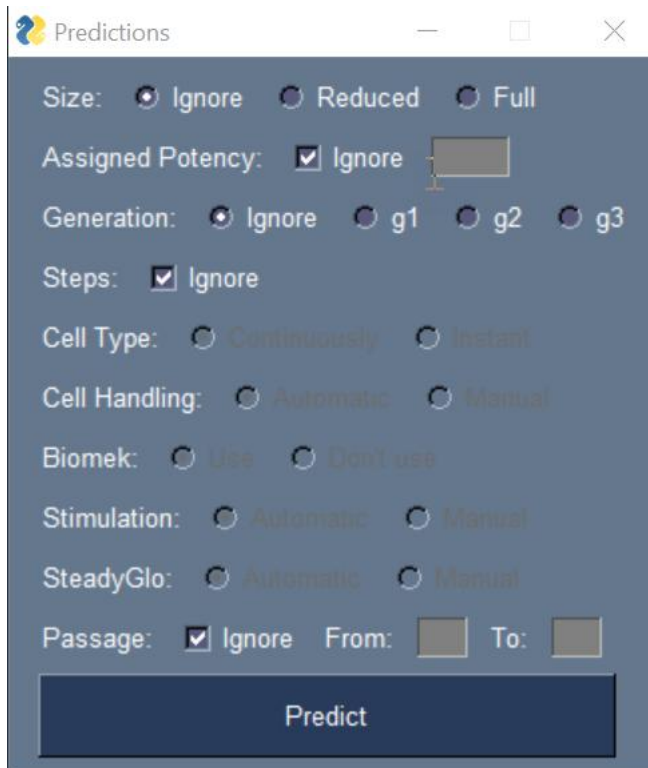
- We have all these data – Use them



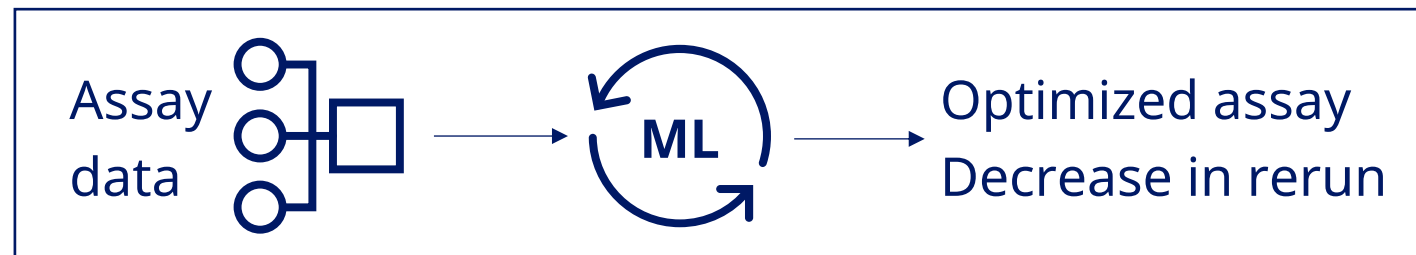
Random Forest ML

- Random Forest is a popular supervised machine learning algorithm that creates an "ensemble" of multiple decision trees, training them on random subsets of data (bagging) and features to reduce overfitting
- $y = ax + b$ used to be the formula for a straight line. Now it is an algorithm

Prediction of assay outcome by ML

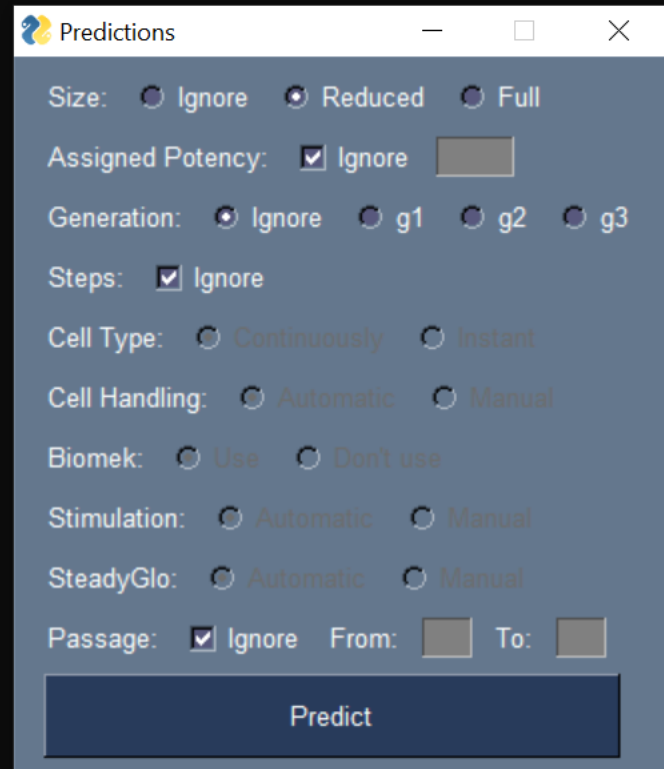


- Predict outcome of an assay, i.e. use certain parameter
- Integrate as part of assay development & performance
 - *Research* ▶ *Phase I/II/III* ▶ *QC*
 - Save time during development
- Useful for any analytical procedure where data exists



Prediction of assay outcome by ML

```
Input:  
Size: Reduced  
Prediction made based on 2267 positions  
Predicted probability of passing potency = 97.8 %  
Predicted probability of passing weight = 93.2 %  
Predicted probability of passing all SST parameters = 89.9 %
```



The screenshot shows a window titled "Predictions" with the following settings:

- Size: Ignore Reduced Full
- Assigned Potency: Ignore
- Generation: Ignore g1 g2 g3
- Steps: Ignore
- Cell Type: Continuously Instant
- Cell Handling: Automatic Manual
- Biomek: Use Don't use
- Stimulation: Automatic Manual
- SteadyGlo: Automatic Manual
- Passage: Ignore From: To:

Conclusion – Data can be used for more than just statistics

- Reduce reruns
- Reduce assay failure
- Standardise assays

- Use of ML
- Predict outcome of an assay based on the different changeable parameters
- The program needs to be automated. Too many parameters to change

- **Do we need a ML program when we have programs like the Power BI?**

The Biggest Challenge

- Some people see the troubles and new rules before they can see the benefits



Acknowledge

- Jon Høj Christensen
- Simon Rueskov Schleicher
- CMC Bioanalysis Team 6
- ChatGPT-5mini

Thank you for your attention

