AI-Accelerated Stability for Biologics

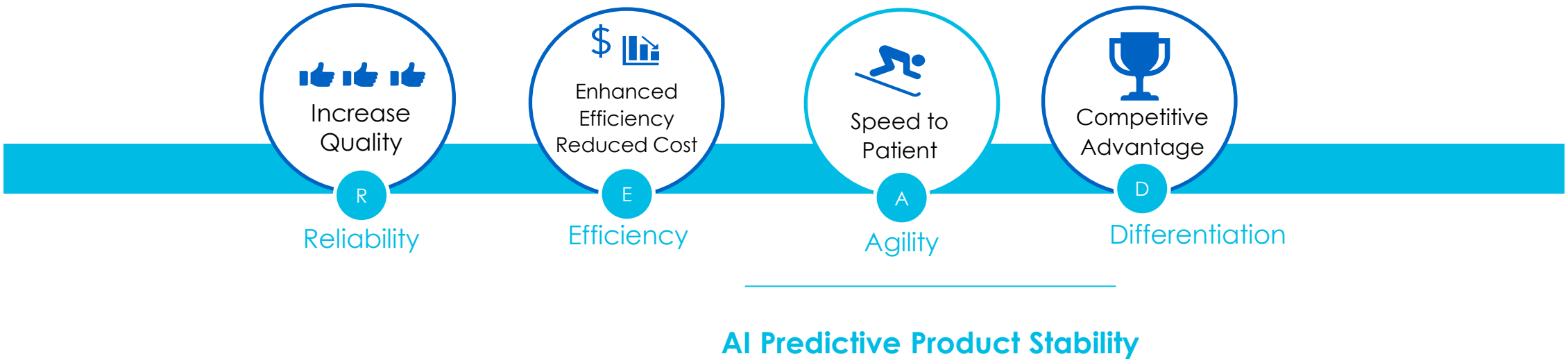
Seyma Bayrak, Director, Global Quality Analytics and Innovation

CASSS CMC Strategy Forum North America, January 27 2025

AMGEN

AI can drive transformational value in every part of our business

The rapid advancements in technology and **AI Platforms** significantly **lowered the barrier of AI-driven innovations and** improve our capabilities to **serve our patients**

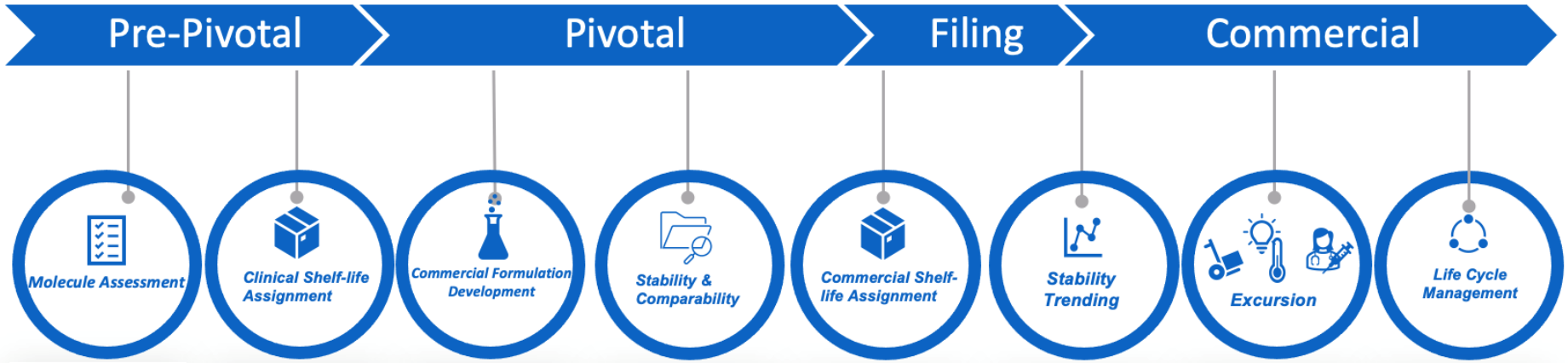


Obtaining long-term stability data is a key rate-limiting step for product advancement

- Need to bring our medicines to patients faster
- Stability remains on critical path in CMC acceleration
- Shelf life assigned based on real-time data presents clinical and commercial challenges

- ✓ ICH Stability Guidance is expected to be restructured & revised to modernize: Expected to include science-based and risk-based stability modeling approaches
- ✓ Significant prior knowledge & data available, advances in understanding of protein structure, product stability and quality attributes
- ✓ Strong data and analytics infrastructure, AI & Data Science expertise

Predictive stability capability can have a favorable impact on the entire value chain:



Predictive Stability Landscape: AI&ML approaches are well suited for stability prediction for biologics

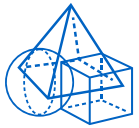
Computational Methods

Data-Driven

Historical data from similar molecules are required

Prior-Knowledge Approach
Linear extrapolation, manual classification of molecules

Machine Learning



First Principles

Understanding of the protein structure and degradation pathways are required

Advanced-Kinetic Modeling

Hybrid Modeling

Combination of first principles and data-driven models when understanding of the mechanistic details is lacking

Small Molecules

- $ASAP_{PRIME}$ is broadly used in the industry
  - It is accepted in development and seen as 'supportive' for commercial shelf-life by regulatory bodies

Vaccines & Biologics

- Advanced kinetic modeling[1] has been successfully used for vaccines (in conjunction with conventional stability submissions)
- Use of prior knowledge for biologics[2]

- **AI/ML models**
  - Ability to learn complex relations/patterns from a large amount of prior knowledge data

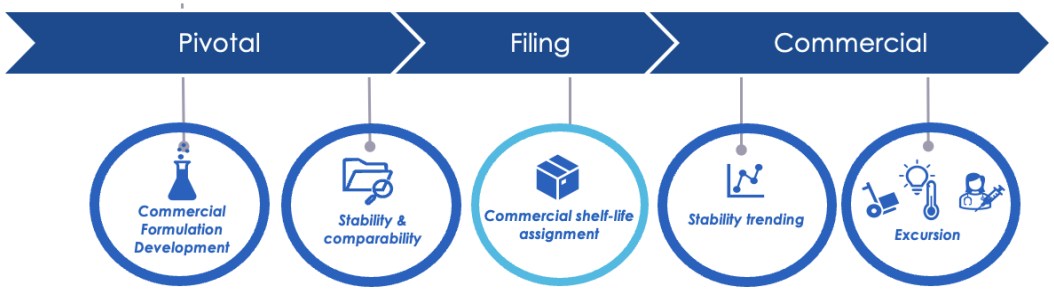
Potential Stability impacting factors for biologics:

- *Modality*
- *Container closure system*
- *Storage conditions*
- *Characterization of Degradation Pathways*
- *The pharmaceutical form*
- *Formulation and Protein Concentration*
- *Stability-indicating assays*
- *Manufacturing processes*
- …

[1] Clenet, *Accurate prediction of vaccine stability under real storage conditions and during temperature excursions J.EJPB* (2018)
[2] Andrew Lennard et al. *Using Prior Knowledge for Stability Modeling of Biological Therapeutic Agents to Assign Shelf Life* (2021)

AMGEN

Case Study: Stability Specification Limit and Commercial Expiry Setting Using Machine Learning for a mAb



Question to be answered

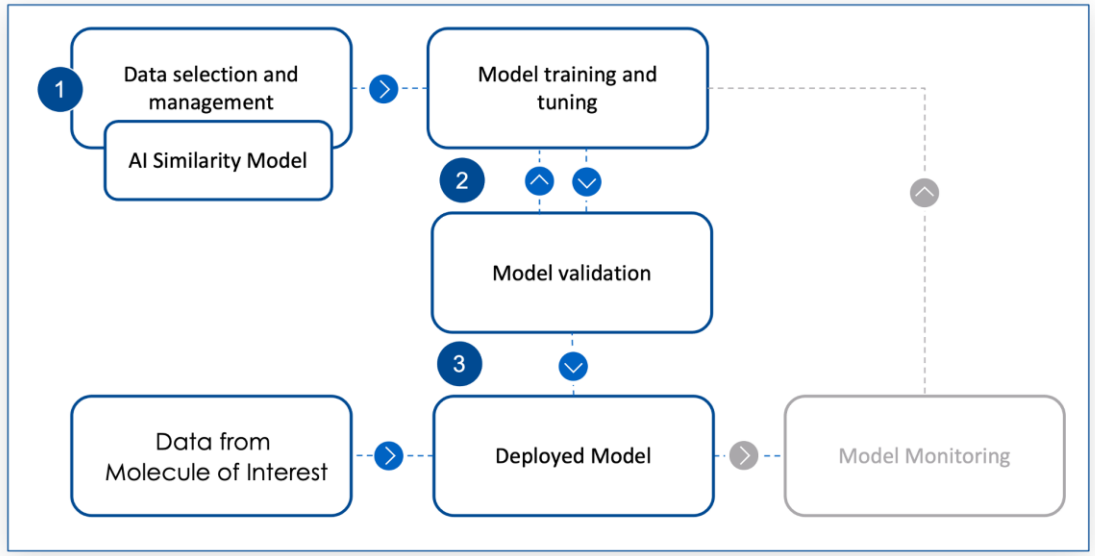
What is the HMW species formation of Molecule of interest at 36-months given 24-months of real-time data at recommended storage conditions (RSC)*

Context of Use

The prediction can be used to support the establishment of stability specifications for the Marketing Authorization Application (MAA).

The final recommendation for the specifications will be provided by Product Quality SMEs.

AI framework is developed following standard best practices in data science any regulatory guidance available to **minimize risk and ensure model credibility**



Proposed Regulatory Framework for modifications to Artificial Intelligence/Machine Learning (AI/ML)-based software as a Medical Device (SaMD) FDA

HMW: High Molecular Weight

Case Study: Stability Specification Limit and Commercial Expiry Setting Using Machine Learning for a mAb

1 Data Selection and Preprocessing

- A rigorous and iterative data selection process was performed **with SMEs** to ensure relevant and accurate data was included in the model
- Only commercial mAbs with liquid drug product were included in the analyses

2 Model Training and Validation

- Multiple machine learning algorithms were evaluated to find the best model that learns the dynamics of the stability profiles (Linear regression, Random forest, Deep Learning etc.)
- **No data from Molecule of interest** was used in the model training or testing process

3 Deployed Model for Molecule of interest

- Developed model was employed to predict the stability of Molecule of interest, predictions were compared with traditional models and **reviewed by SMEs**

We have built an AI model to find the most similar mAbs to Molecule of interest

1 Data selection and management

AI Similarity Model

Based on %HMW variations, model suggested mAb1, mAb5 and mAb3 are more similar to molecule of interest

Data used for Similarity Modeling

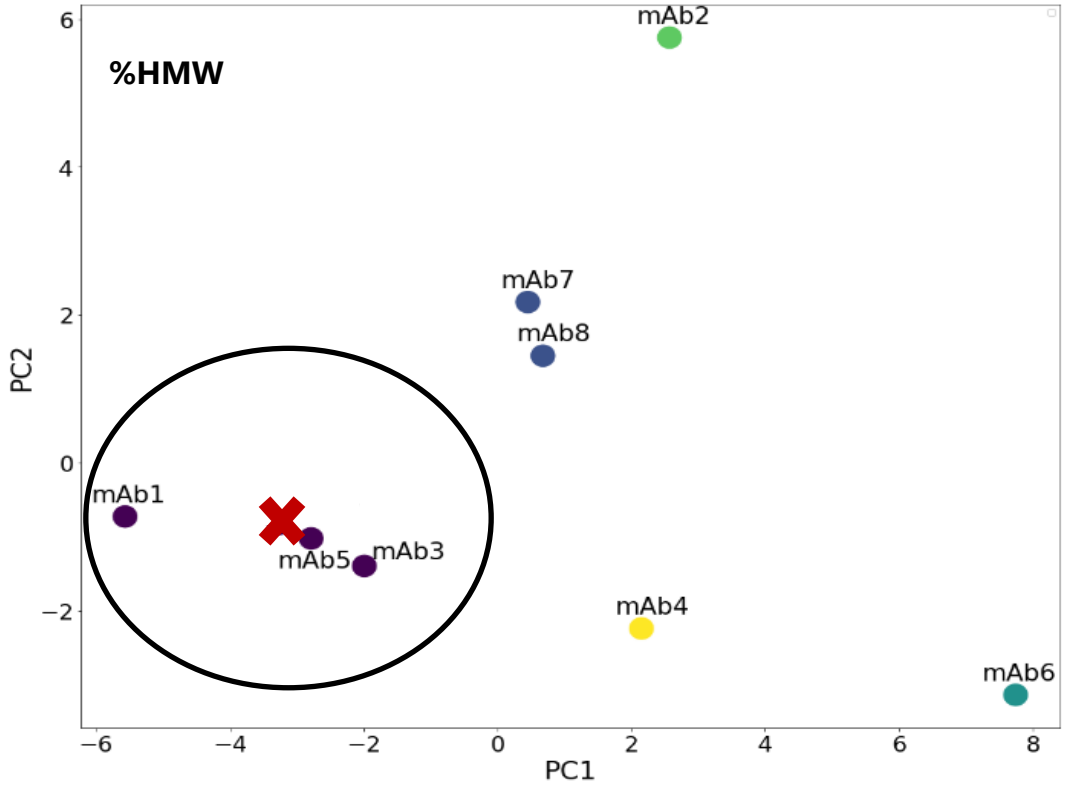
Data of up to 24 months were used for %HMW and other potential stability-impacting attributes to find the most similar historical molecules to molecule of interest

Multiple combinations of features are tested to make sure the results were consistent (lots are summarized using statistical features to find product level similarity (mean, std, max, min, percentiles, etc.):

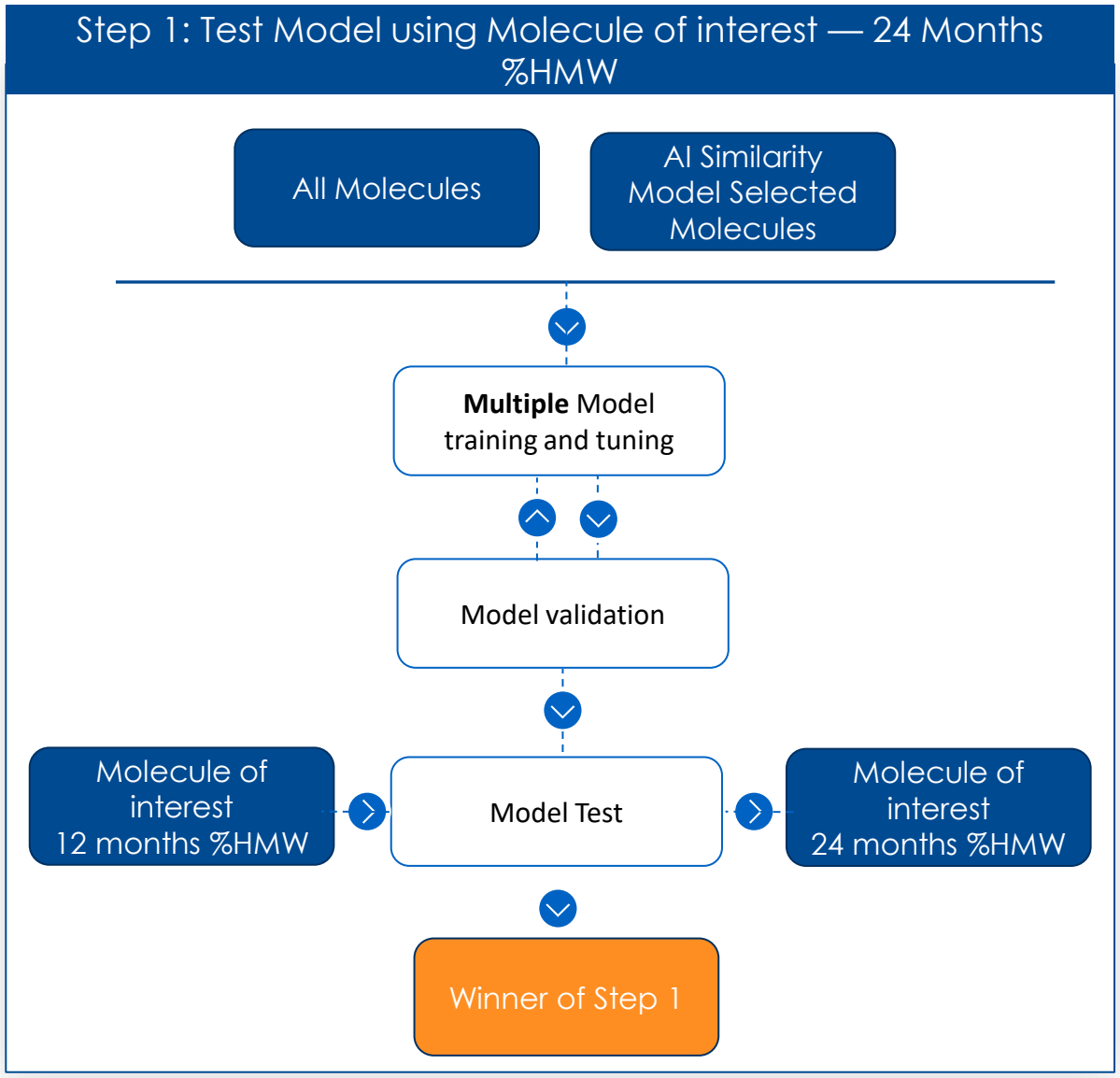
- Only %HMWs
- %HMWs + pH
- %HMWs + pH + Concentration

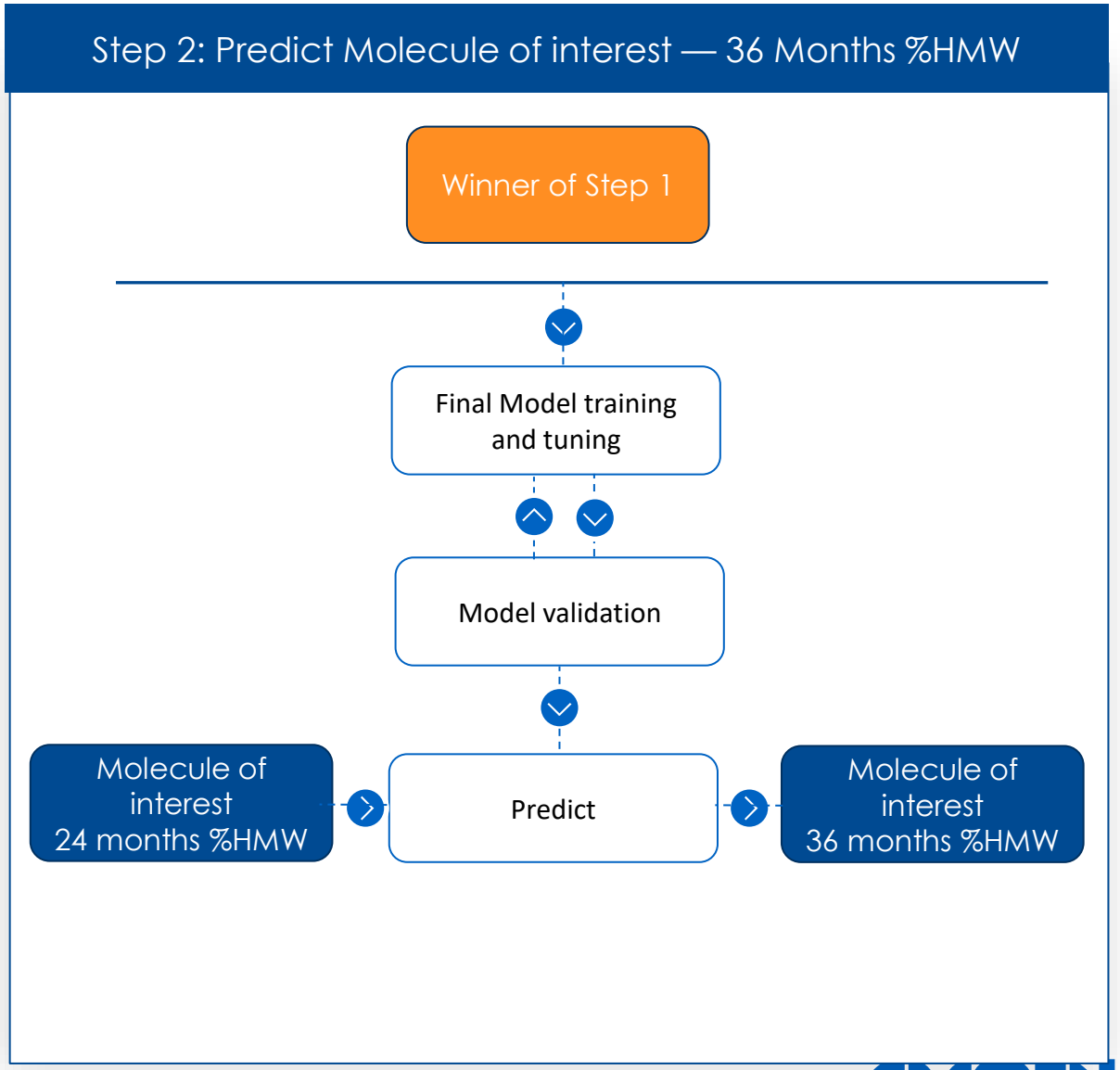
Principal Component Analysis (PCA) is performed prior to applying an unsupervised Mean Shift clustering method

%HMW based similarity



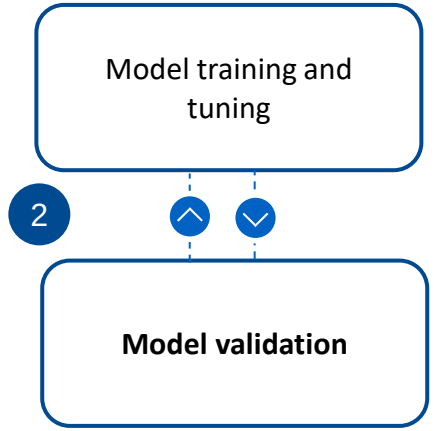
Two models have been developed using different combinations of data sets and two step approach was taken to evaluate the model performance

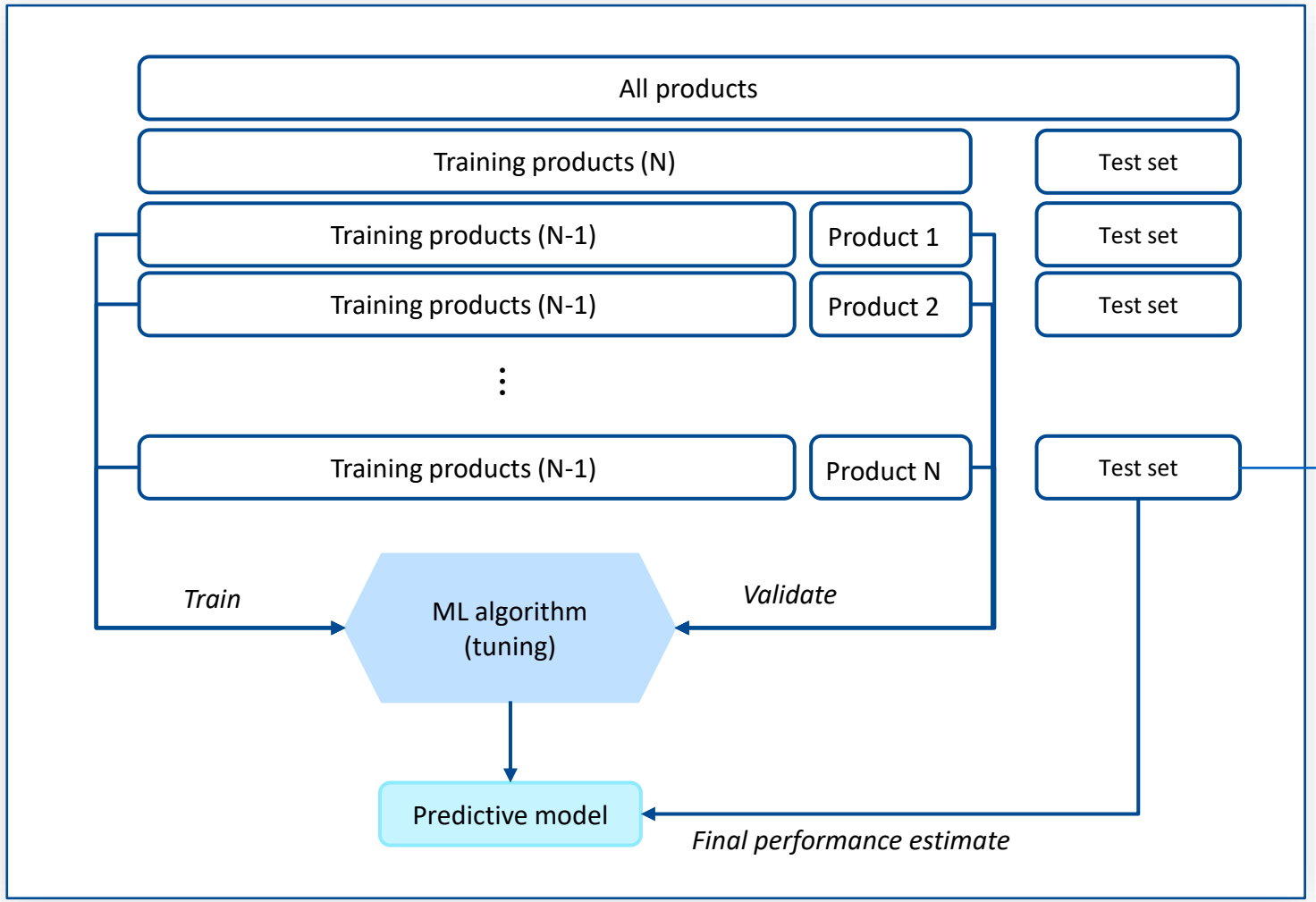




Standard best practices are followed to ensure the accuracy and reproducibility of developed models

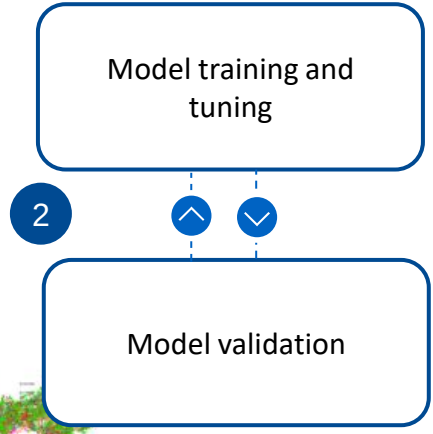
We adopted the leave-one-out (LOO) approach, a customized version of the k-fold cross-validation to ensure models are not overfitting but generalized well for products they have not seen.





Test product was selected based on the similarity analysis and was never used in the training process

Neural Network Algorithm was Selected based on the model performance of Step 1



Various algorithms were initially tested as prototype models for prediction accuracy using a test product

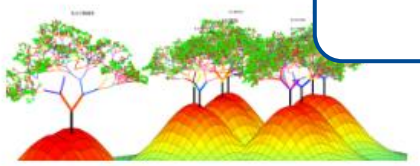
- Selected a neural network algorithm

The structure of a neural network model is defined by hyperparameters:

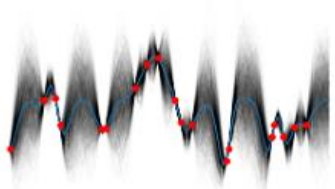
- Hidden layers size
- Number of hidden layers
- Batch size
- Loss function
- Early stopping criteria

} Values determined based on iterative training

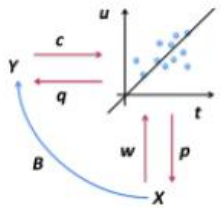
Artificial Neural Networks



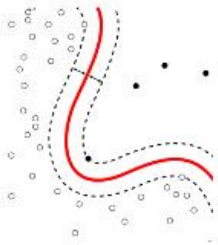
Random forest



Gaussian Processes

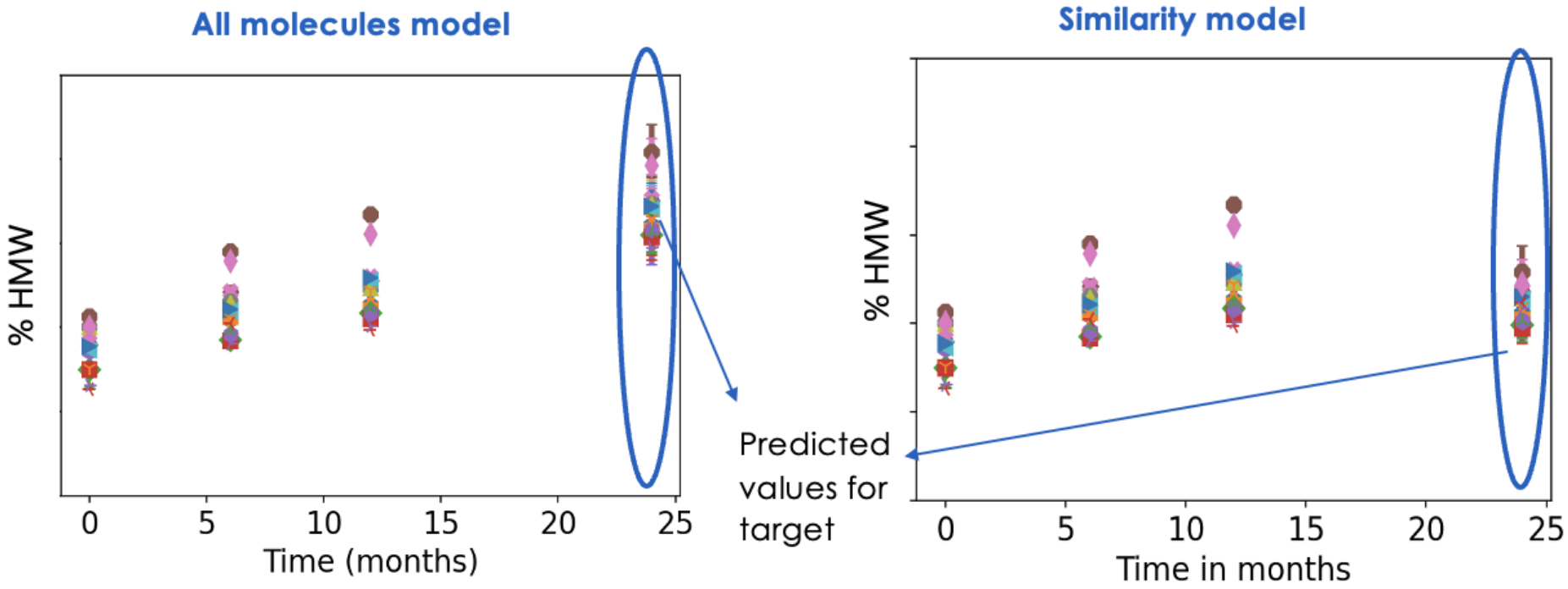


Partial Least Square



Support vector machines

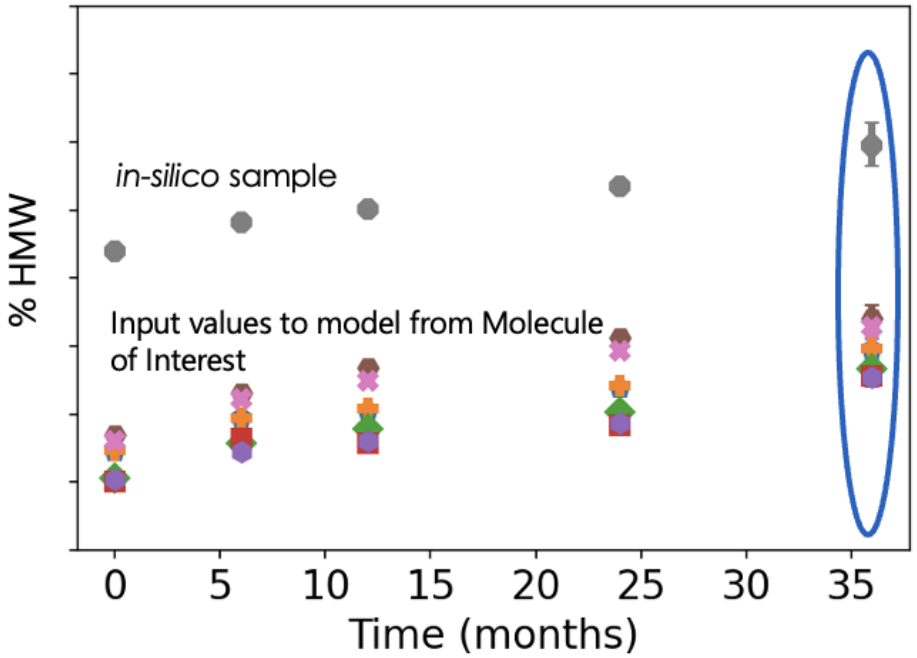
Results- Step 1 *24-month predictions: Model based on all prior knowledge molecules has performed better*



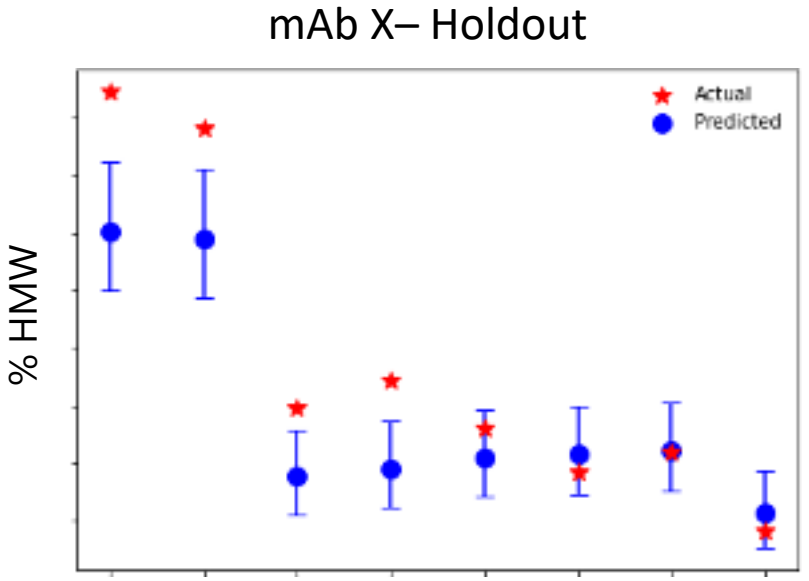
- Initial and/or prior HMW values predominantly determine the long time HMW values
- Besides HMW values, concentration and pH also had measurable impacts
- Predictions from ML model built using all the prior knowledge molecules (highly similar and less similar) are more accurate

Results – step 2: *Model based on all prior knowledge molecules was selected as final approach and used to predict %HMW at 36 month*

3 Deployed Model



Predicted values



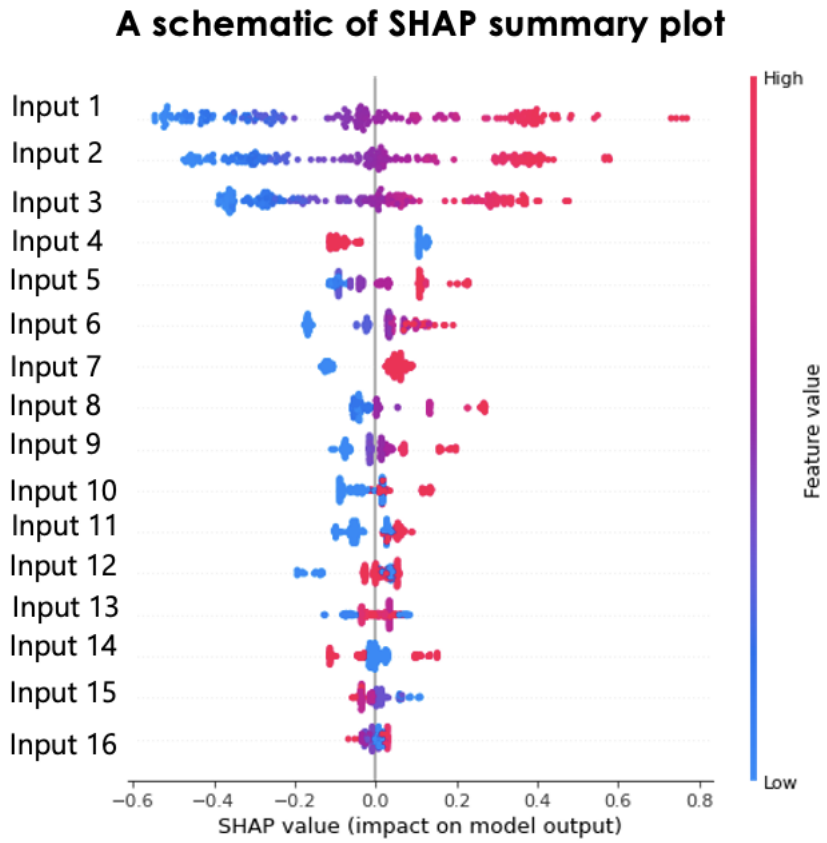
Generalized Error	~9.2%
mAb X– Holdout	~9.4%

- Model predictions in line with the estimates from conventional methods (e.g., common slope estimate)
- Predictions help in setting up the stability specifications for the drug product

In silico Sample is generated to represent the worst-case scenario of HMW at the upper release limit at time = 0

AMGEN

Impact of Features on %HMW predictions using SHAPley Analysis



x-axis: SHAP value

y-axis: input variable names, highest to lowest impact (*T -> B*)

Number of points for each input variable: number of training + target samples

Positive correlation: blue to red (*L -> R*)

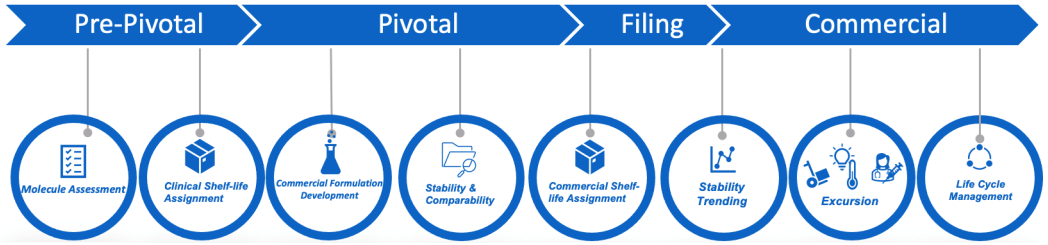
Negative correlation: red to blue (*L -> R*)

Spread of the points: broader (higher impact), smaller (lower impact)

T: Top, B: Bottom, L: Left, R: Right

Conclusion

Predictive Stability enables many opportunities to optimize, accelerate and improve across the value chain



Predictive stability models present a differentiating capability leads to rapid development of timelines to bring our medicines to clinics and market faster

When built effectively they can reduce stability risk,
help ensure quality and supply

Predictive stability - a cross-functional effort with members from Amgen GRAAS, GQ, OT&DS, and PD

Special mention

Dr. Sriramvignesh Mani
Dr. Andrew Lennard
Dr. Camilla Santos
Dr. Gerd Kleeman
Frank Pagliusi
Saleh Alkhalifa
*Dr. Barbara Rellahan**

GRAAS: Global Regulatory Affairs and Strategy
GQ: Global Quality
OT&DS: Operations Transformation and Digital Strategy
PD: Process Development

* ex-Amgen staffs