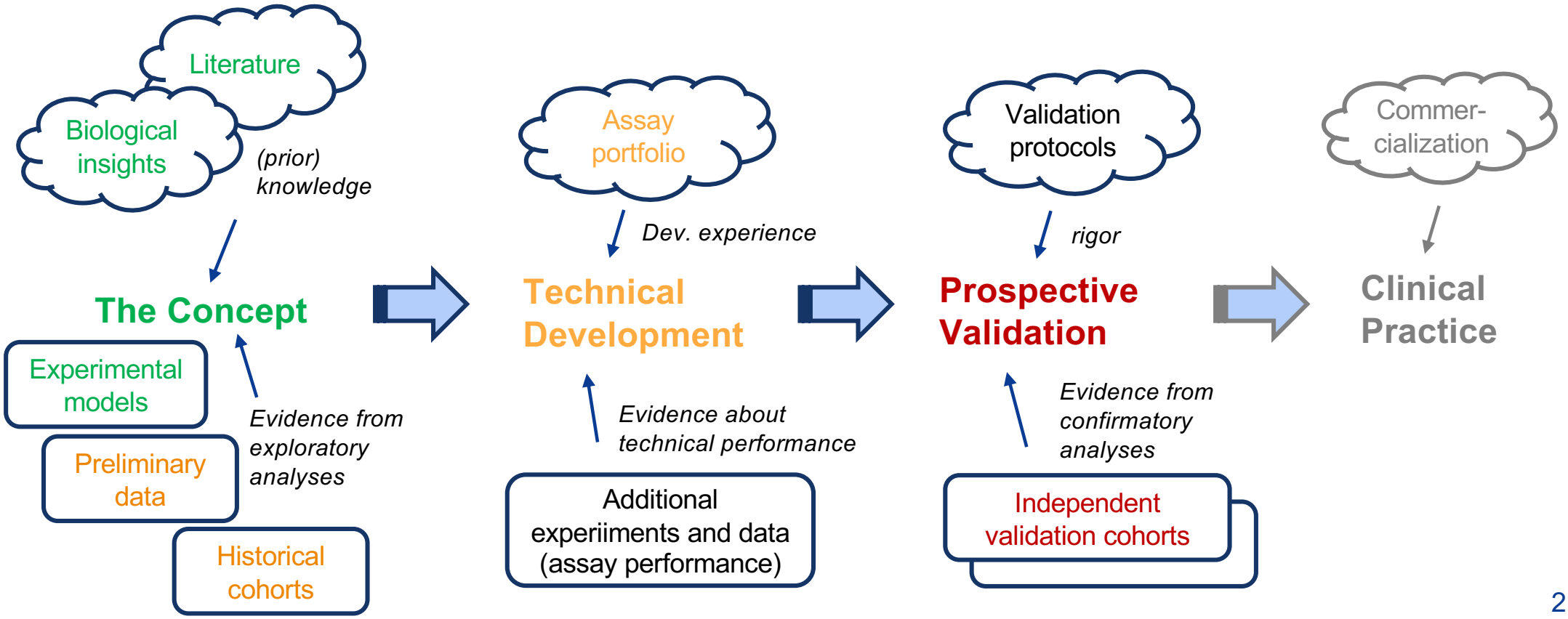
# Reproducibility of Evidence Generation in Biomarker Development

Anton Belousov, Roche

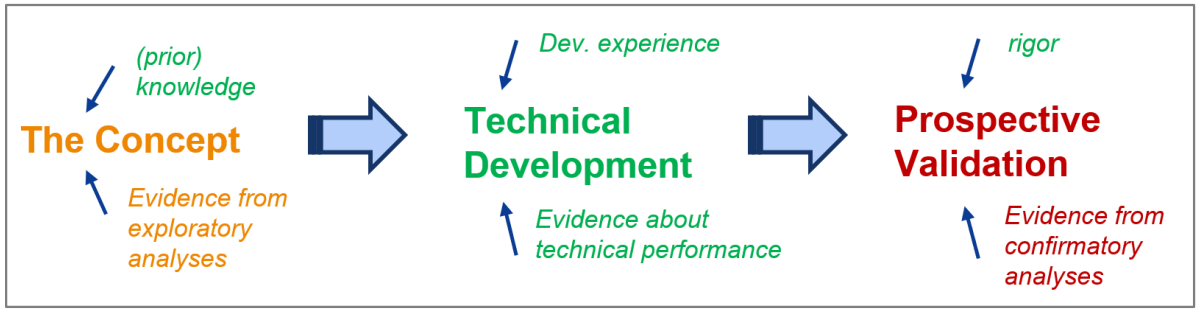
BBS Seminar, April 2024

# (Biomarker) Development as Evidence Generation Process. A Chain is not Stronger than its Weakest Link

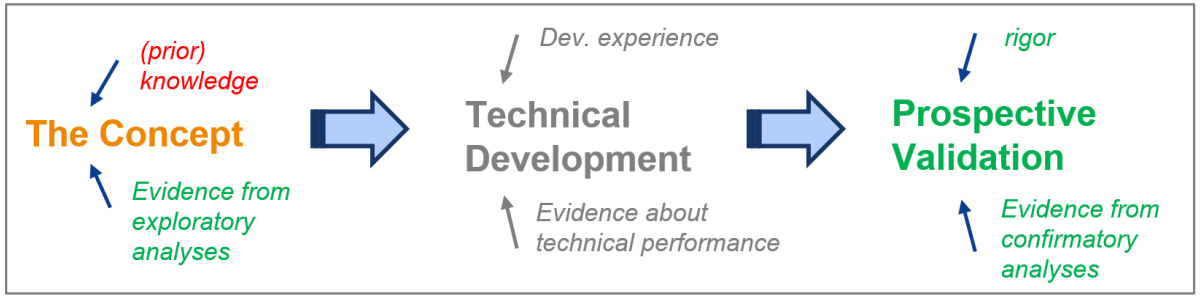


# Two Illustrative Examples

1. **Biomarker predictive of response to Cibisatamab CEA TCB (T-cell Bispecific)**




2. **Model (Exposure + Biomarker) to predict risk of CRS in PTs treated with Columvi CD20-TCB**



# Two Illustrative Examples

## 1. Biomarker predictive of response to Cibisatamab CEA-TCB (T-cell Bispecific)

- Anticipated clinical practice: Screening to increase probability of response, MSS CRC
- Assay format: Multiplex rtPCR (gene expression) on FFPE biopsy

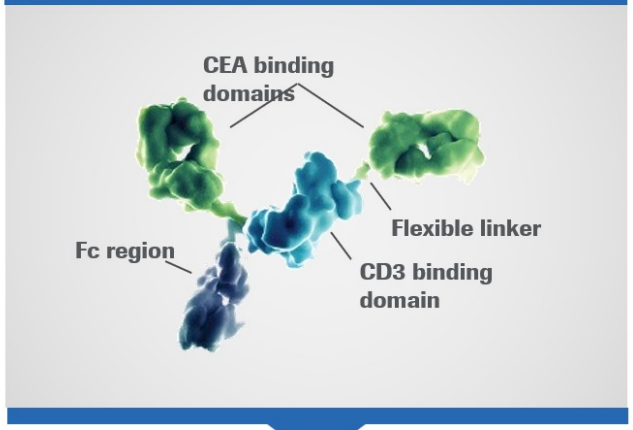
## 2. Model (Exposure + Biomarker) to predict risk of CRS in PTs treated with Columvi CD20-TCB

- Anticipated clinical practice: Baseline preciction of the likelihood of Gr2+ CRS after the first dose of Columvi (NHL)
- Assay format: Multi Analyte Algorithmic Assay (clinical; central lab; radiology)

# Cibisatamab, a CEA-CD3 TCB

## Antibody Design

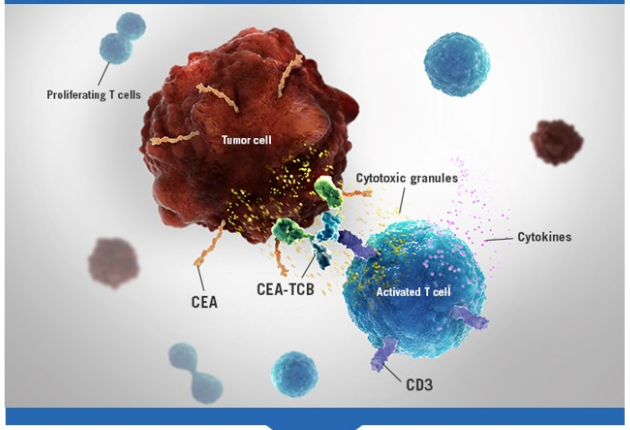
**2:1 CEA/CD3 TARGETING**




- Targeted T-cell activation against CEA-expressing tumors
- Tight T cell-tumor cell connection
- Extended half-life

## MOA as Monotherapy

**ENGAGE and ACTIVATE**




- Simultaneous binding to T Cells and tumor cells
- Activation of T cells for potent, targeted tumor cell killing
- Recruitment of additional T cells

References: Bacac, M. Clin Cancer Res. 2016; 22(13), DOI: 10.1158/1078-0432.CCR-15-1696
Tabernero, J. JCO. 2018; 35(15), DOI: 10.1200/JCO.2017.35.15

# Biomaker Hypotheses and Supportive Data

**CRC MSS**

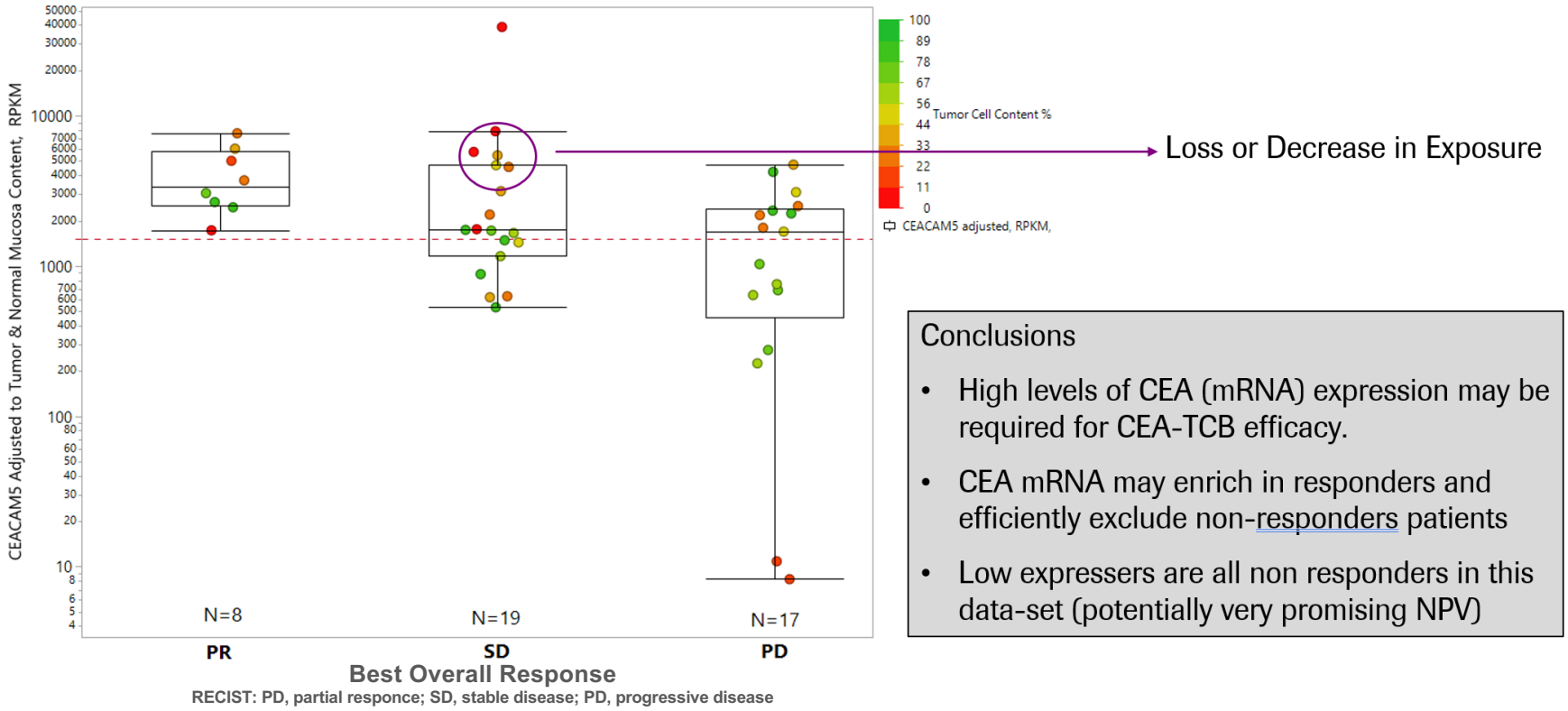
| | |
|---|---|
| CEA-TCB 160mg flat dose QW | N=37 |
| CEA-TCB 100mg flat dose QW | N=20 |
| CEA-TCB 100mg flat dose Q3W | N=19 |

**Confounding factors:**
Loss of exposure, Immunogenicity, Safety

| Hypothesis | PD BM insights |
|---|---|
| CEA-TCB induces an increase in intratumoral immune cells | 1 Treatment with CEA-TCB increases T-cell infiltration (activated CD8 T cells and PD1 levels) in tumors |
| CEA-TCB targets cells that express CEA | 2 Treatment with CEA-TCB could lead to killing of only CEA high cells |
| CEA-TCB requires a certain threshold of CEA in the tumor to trigger anti-tumor activity. | 3 Tumor CEA levels appear to predict for response |

# CEACAM5 gene expression appears to predict for response

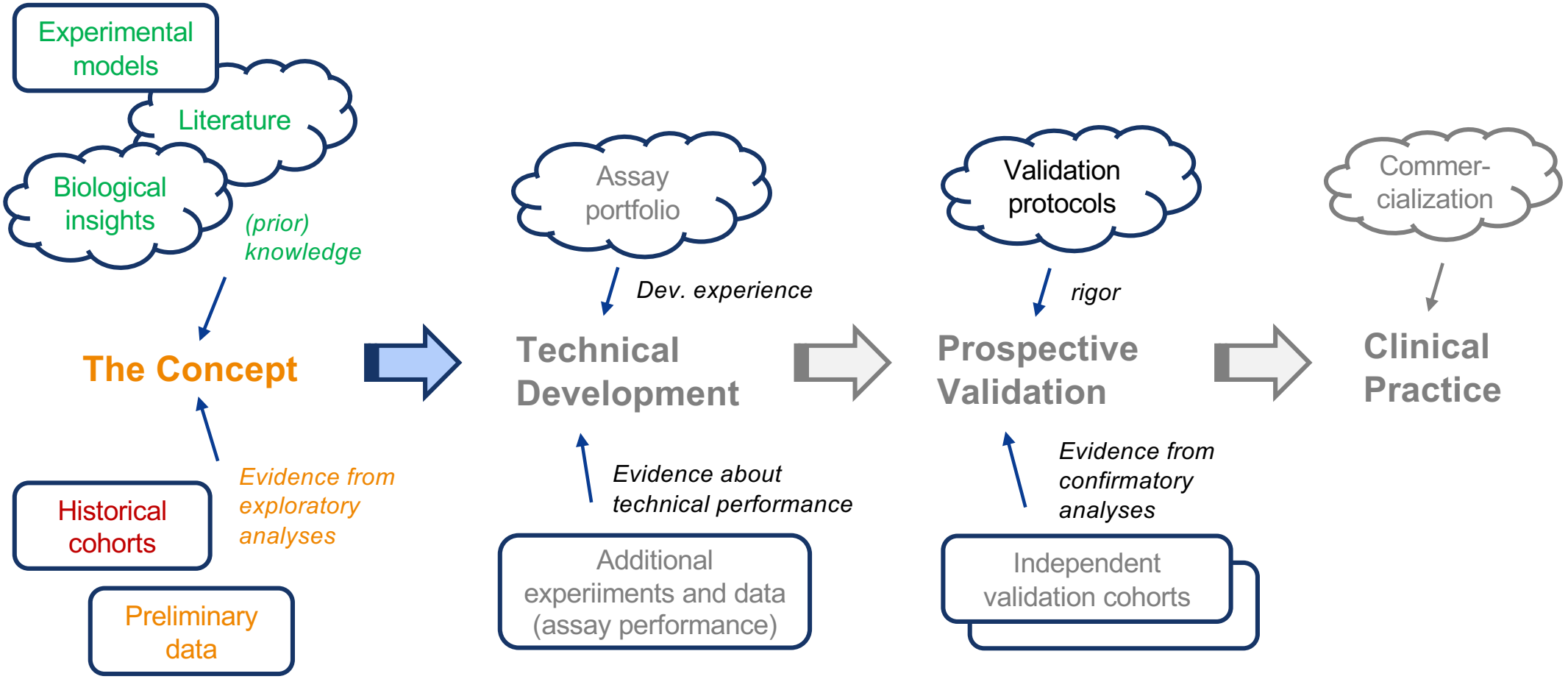
- CEA levels from IHC & RNAseq.
- Analysis of temporal pattern (visit) and expression in archival / fresh samples
- Potential role of study covariates on CEA expression. Primary tumours vs. metastases
- CEA expression in samples of different composition (tumour content, % necrotic tissue)
- Adjusting for prognostic factors and potential confounders (Stage, MSS/MSI status, Mut. load, LOE)




Conclusions

- High levels of CEA (mRNA) expression may be required for CEA-TCB efficacy.
- CEA mRNA may enrich in responders and efficiently exclude non-responders patients
- Low expressers are all non responders in this data-set (potentially very promising NPV)

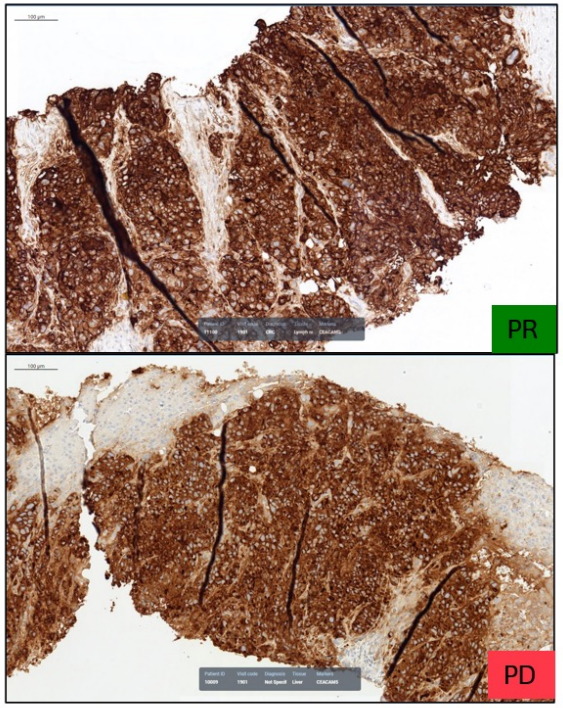
# The Concept: A Victory of Prior over Likelihood…

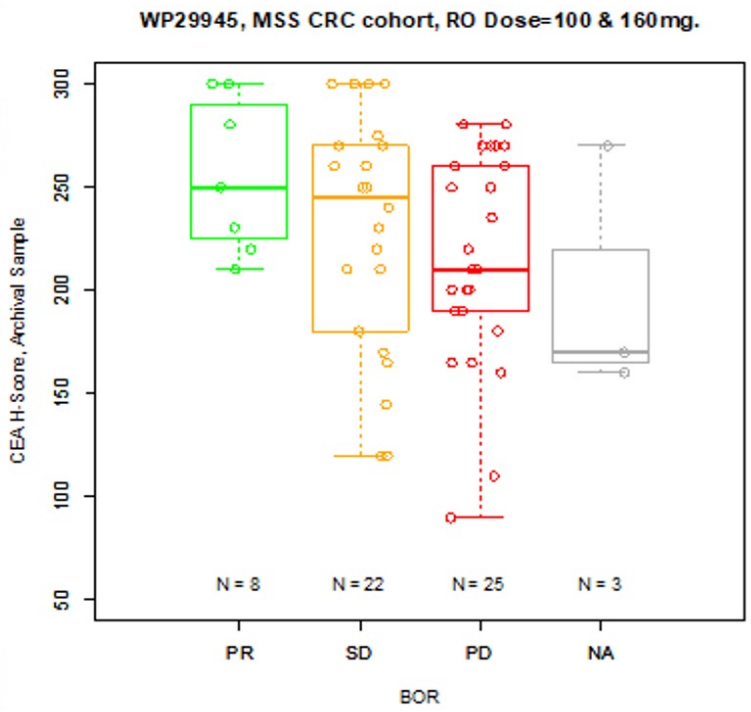


# Predictive Biomarker Needs the Right Assay

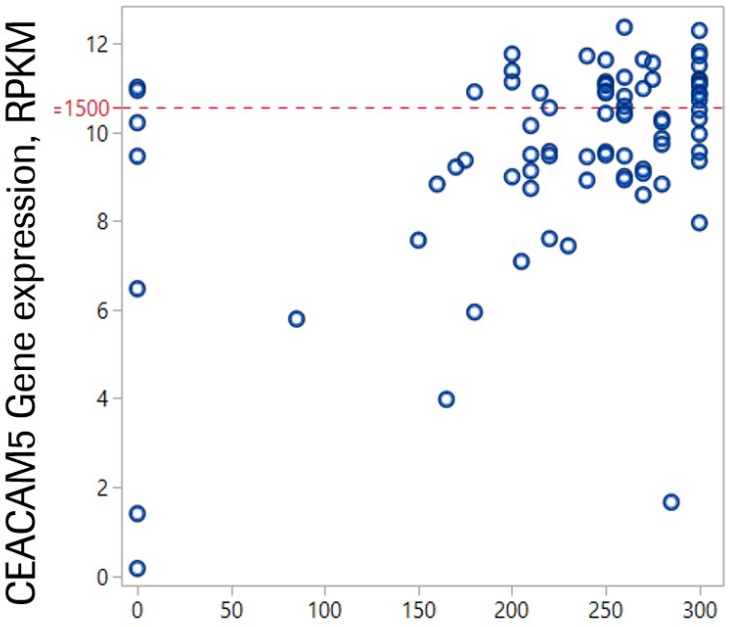
IHC assay for CEA hit the limit of dynamic range: CEA levels by IHC is not Predictive

CEA IHC



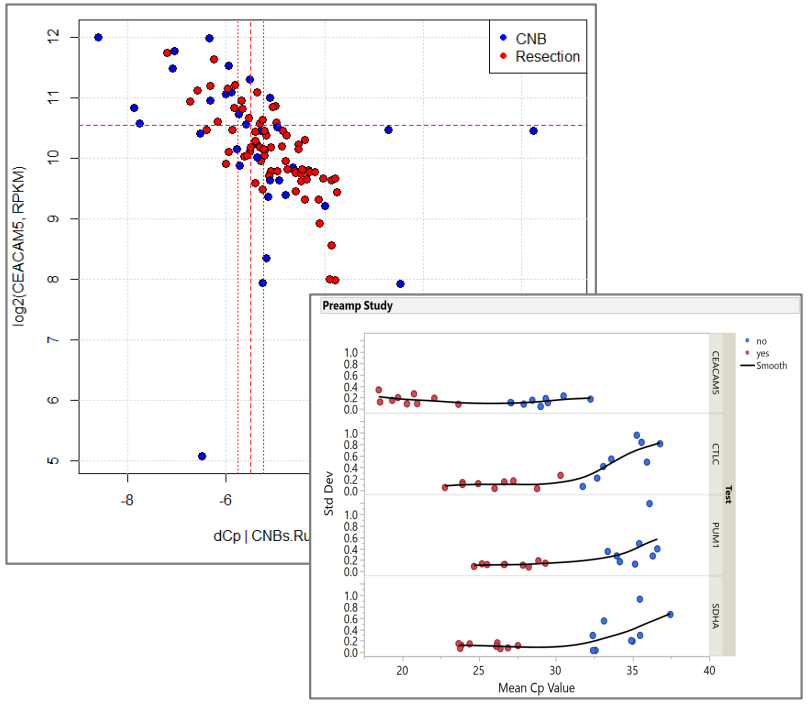

Poor Correlation between IHC and GE



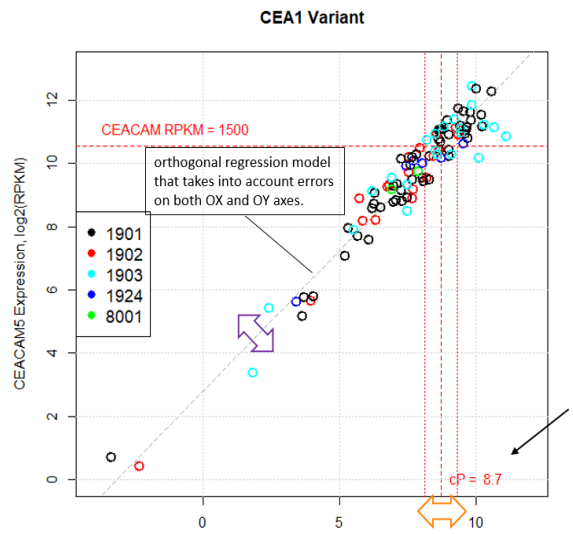

CEA H-score by IHC

# Development of a Multiplex qRT-PCR Assay

- Selection of housekeeping gene panel on historical & external cohorts (RNAseq and Nanostring).
- Clinical samples re-run with 3 designed Thermo Fisher probes (Almac).
- QC workflow and assay design finalized on external samples from Almac (N~100).
- Prevalence study and cutoff selection.
- First technical validation on the clinical samples (N~50, qRT-PCR).
- Second technical validation on a cohort of separately acquired samples (N~125). qRT-PCR and RNAseq.




## Correlation of rtPCR and RNASeq Readouts




- Very good overall correlation
- Linear relationship
- At the 'target' threshold level (RPKM=1500): the variability of cP values that correspond to same RPKM is SD(cP) ~ 0.4 cP unit (1.3-fold)

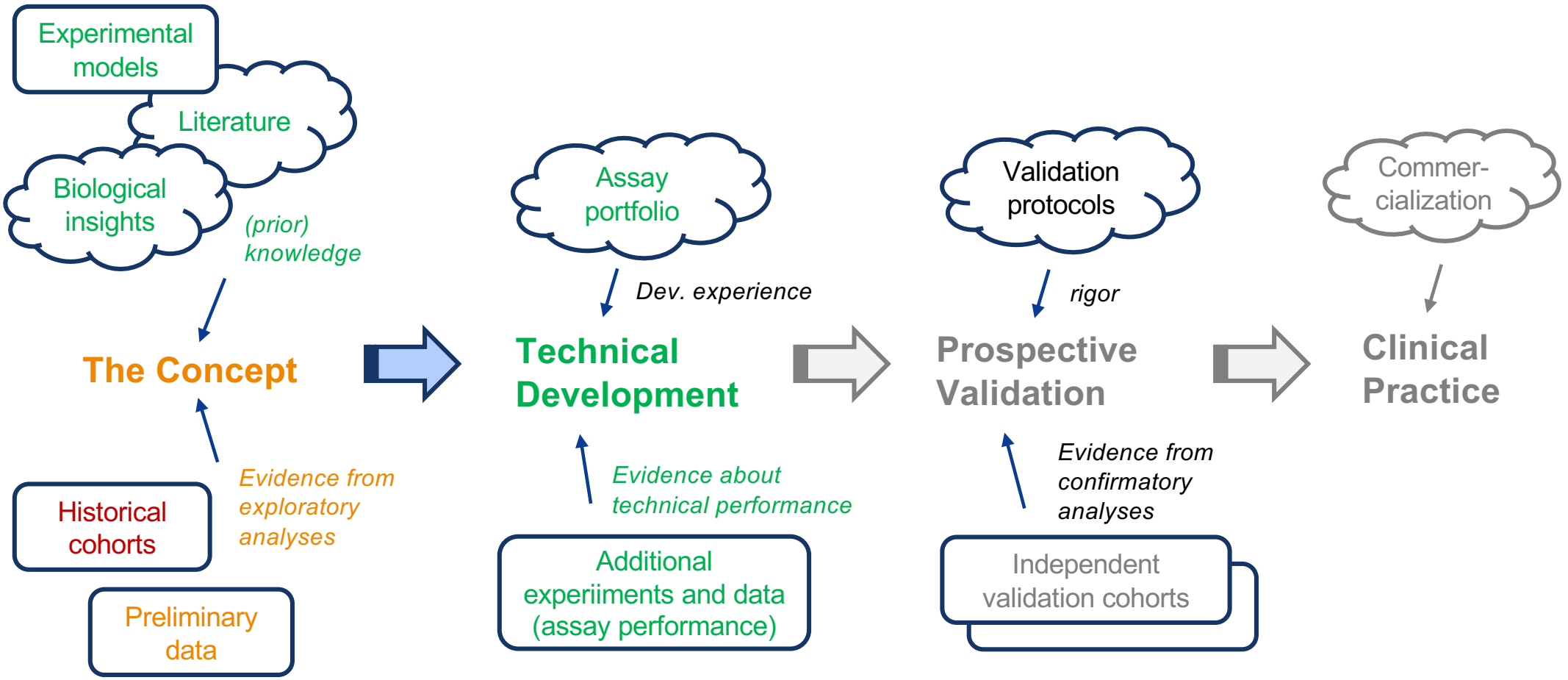
The **threshold** on rt-PCR (cP values) that corresponds to RPKM=1500. Vertical dotted lines: estimated ±2σ uncertainty interval for cP

SD.cP = variability in cP values at around the cP 'cut-off' point (cP=8.7) that corresponfds to RPKM cut-off of 1500.

All available samples independent on RNASeq QC status.

7

# Techical Development: Solid Performance



# Phase 1b/2 Study to Optimize Benefit-Risk Profile

**Stage IV MSS CRC**
- ≥ 2 prior CT regimens
- ECOG PS 0-1
- CEACAM5 high

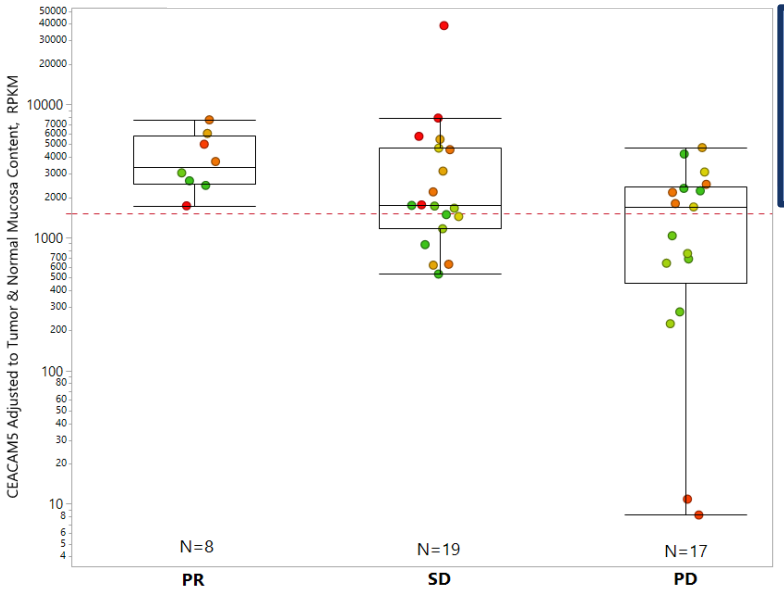
**Safety Run-in**
GpT + CEA-TCB + Atezo
N=6

**Exploratory Phase (Part 1)**
GpT + CEA-TCB + Atezo
N=40

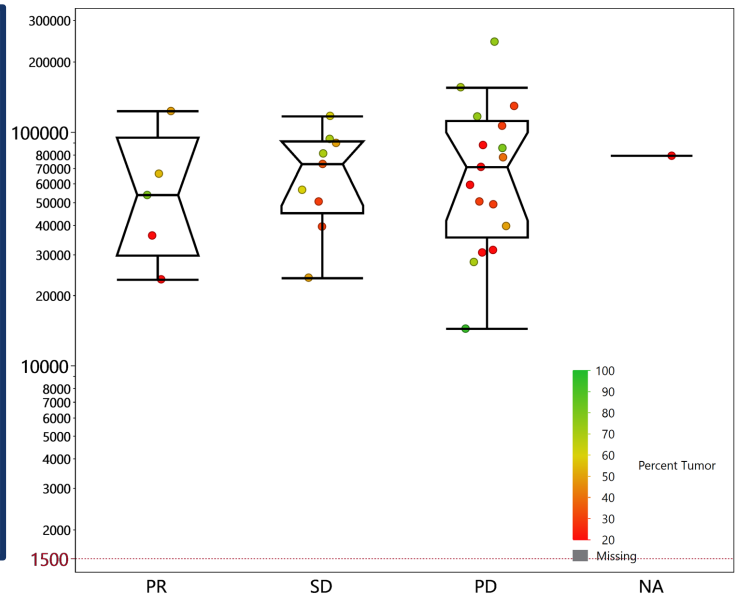
**Efficacy/Safety Gate**
40 pts ORR

- **Biomarker selection** (CEACAM5 gene expression) to maximize ORR
- **Gazyva pretreatment** (2000 mg –d13/12) to mitigate ADAs

**The Concept:**




**Prospective Validation:**




BOR, Best Overall Response (RECIST): PD, partial responce; SD, stable disease; PD, progressive disease

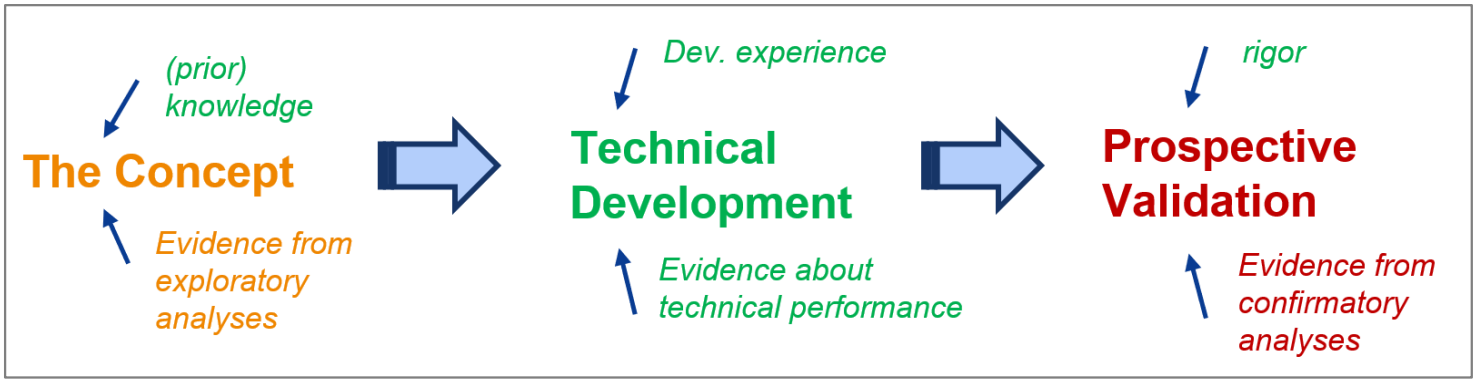
# Biomarker Predictive of Response to Cibisatamab

- CEACAM5 gene expression was identified as a candidate biomarker that may further increase ORR, optimize benefit-risk and enable accelerated clinical development.
- A screening IVD assay (multiplex qRT-PCR on FFPE biopsy) successfully developed and utilized in a Ph 1b trial
- Biomarker does not seem to hold predictive value

**Insufficient initial evidence:**

→ Low N

→ complex interrelation of confounders?



# Two Illustrative Examples

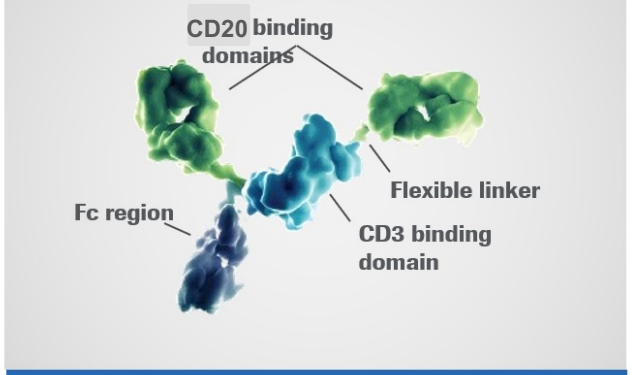
1. **Biomarker predictive of response to Cibisatamab CEA-TCB (T-cell Bispecific)**
   - Anticipated clinical practice: Screening to increase probability of response, MSS CRC
   - Assay format: Multiplex rtPCR (gene expression) on FFPE biopsy

2. **Model (Exposure + Biomarker) to predict risk of CRS in PTs treated with Columvi CD20-TCB**
   - Anticipated clinical practice: Baseline preciction of the likelihood of Gr2+ CRS after the first dose of Columvi (NHL)
   - Assay format: Multi Analyte Algorithmic Assay (clinical; central lab; radiology)

# Columvi, a CD20-CD3 TCB

## Antibody Design

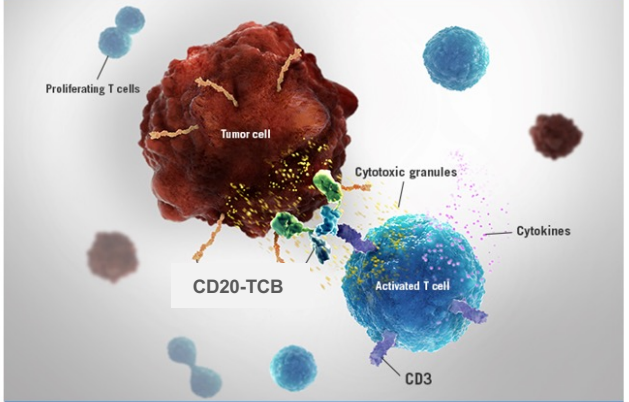
**2:1 CD20/CD3 TARGETING**




- Targeted T-cell activation against CEA-expressing tumors
- Tight T cell-tumor cell connection
- Extended half-life

## MOA as Monotherapy

**ENGAGE and ACTIVATE**

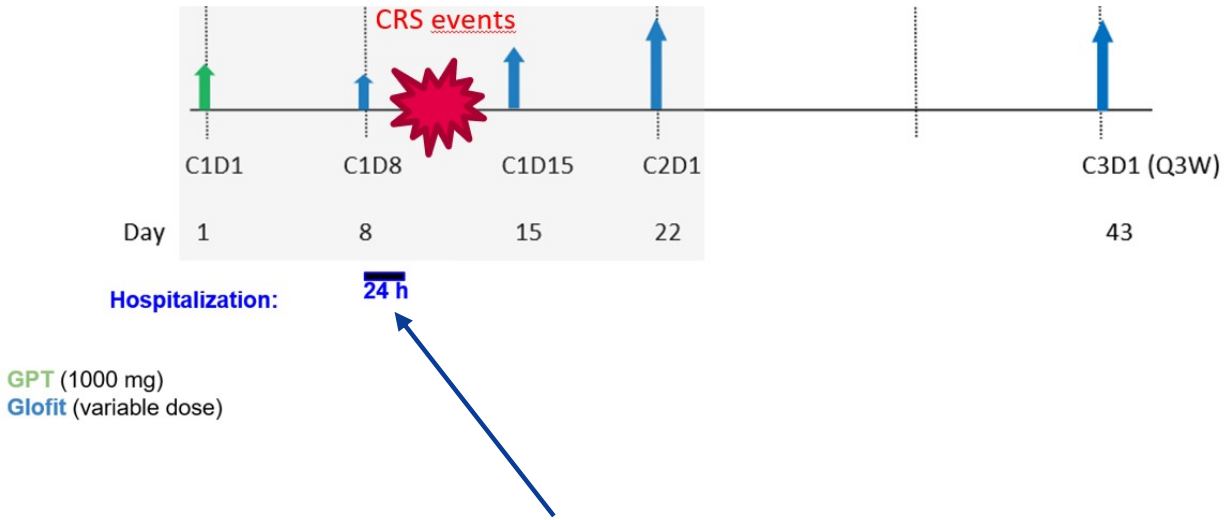



- Simultaneous binding to T Cells and tumor cells
- Activation of T cells for potent, targeted tumor cell killing
- Recruitment of additional T cells

References: Tian et al. J Hematol Oncol (2021) 14:75

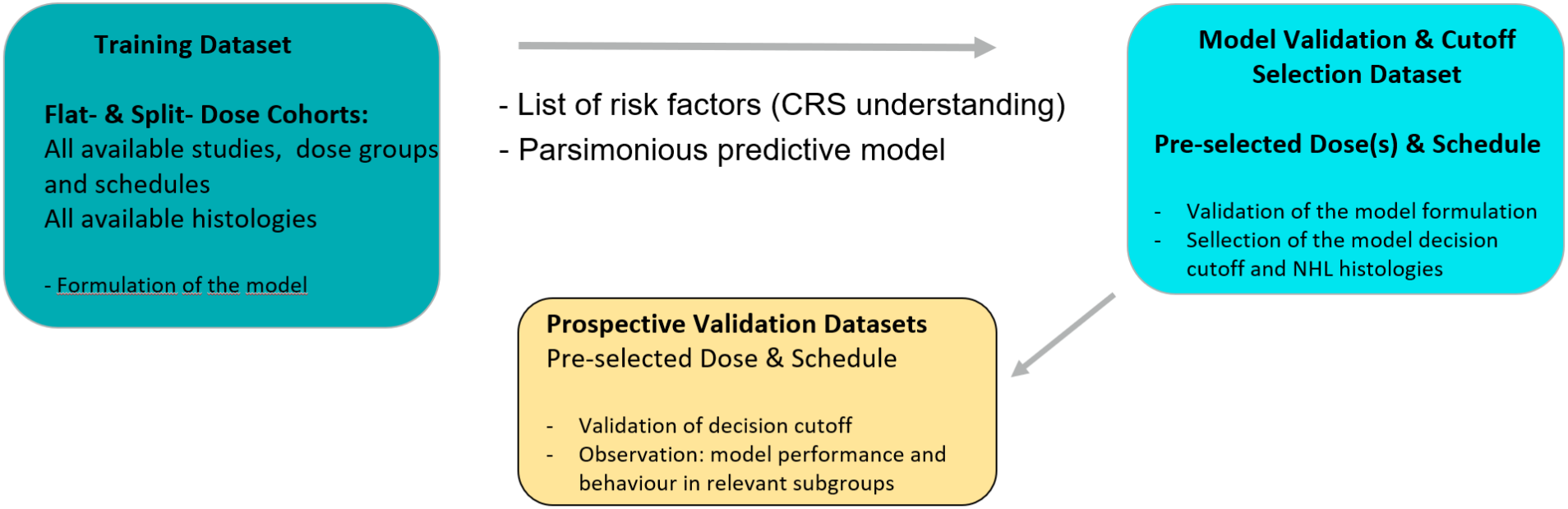
# Need to Predict Cytokine Release Syndrome (CRS)

Identify a meaningful size patient subset (at least 20-25%) with high (>90%) likelihood of staying free of Grade 2+ CRS, for whom outpatient monitoring would be appropriate




Application of CRS-RS may inform investigator decision in clinical trials in future to wave hospitalization requirement for “low risk” patients

# Strategy of Model Development and Validation



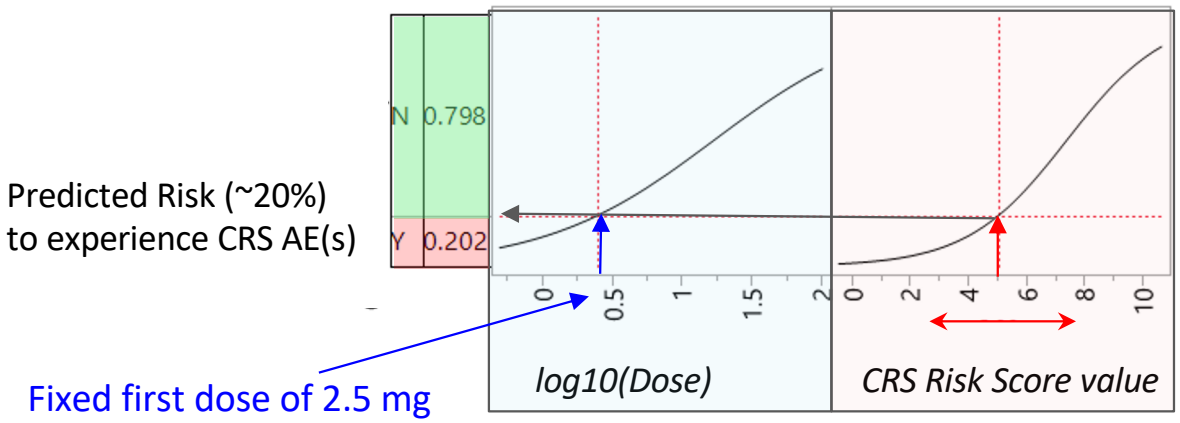

Due to study design limitations and data availability, the model may need to be developed based on the data across several studies and dose groups / regiments.

- Lack of randomization or stratification → potentially strong confounding & mis-balance of many characteristics
- Strong effect of Dose on CRS to be considered → any room for the predictive model to add value?
- Several prognostic/predictive factors may be known → any room for the predictive model to add value?

# The Predictive Model

- CRS-RS: CRS Risk Score is introduced as a weighted sum of (binarized → 0/1) risk factor conditions
- The 8-parameter score can be reduced to a 5parameter score CRS-RS.5p for aggressive NHL histologies
- The final predictive model combines the CRS-RS & Columvi dose to estimate the expected probability of CRS event

***log( Odds Gr2+ CRS ) ~ log(Dose) + CRS Risk Score***

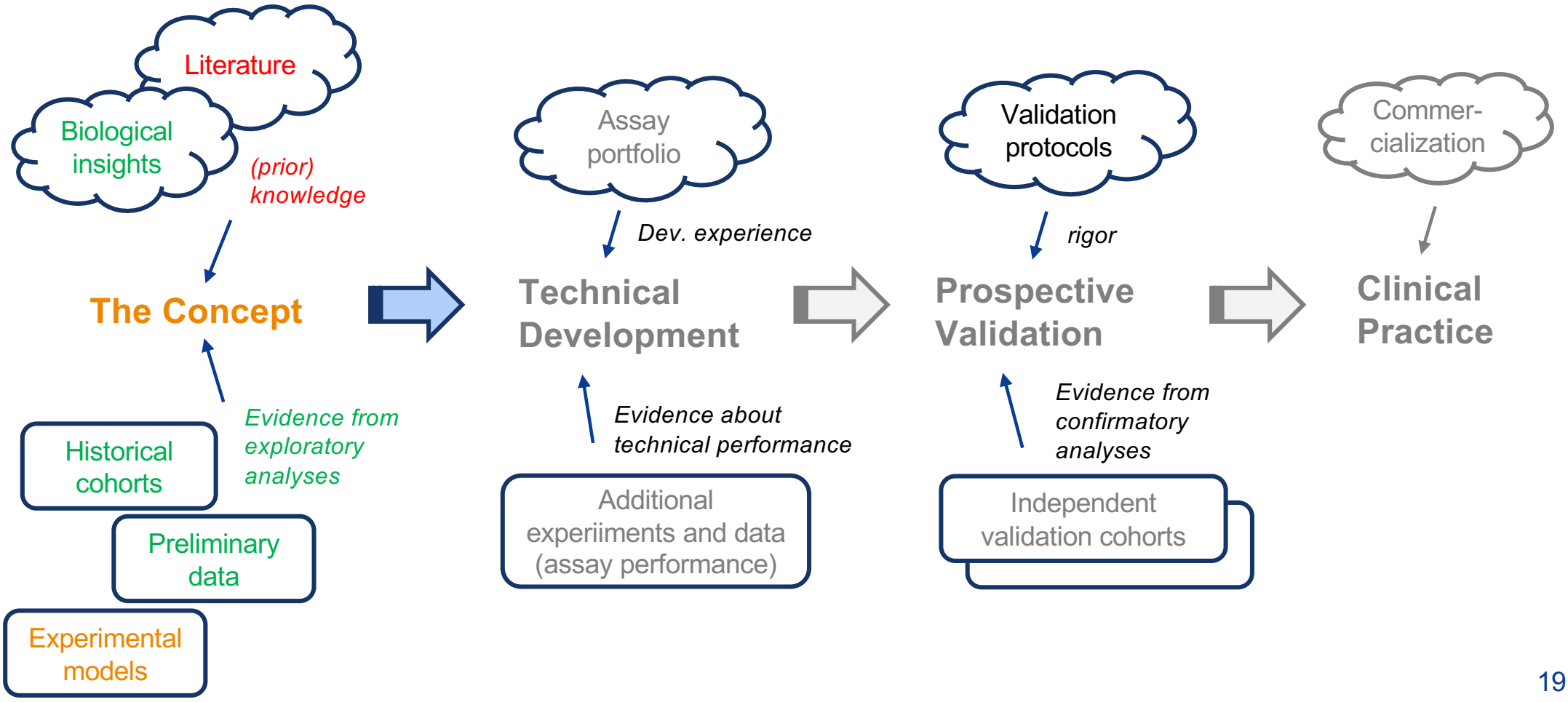



**CRS-RS**

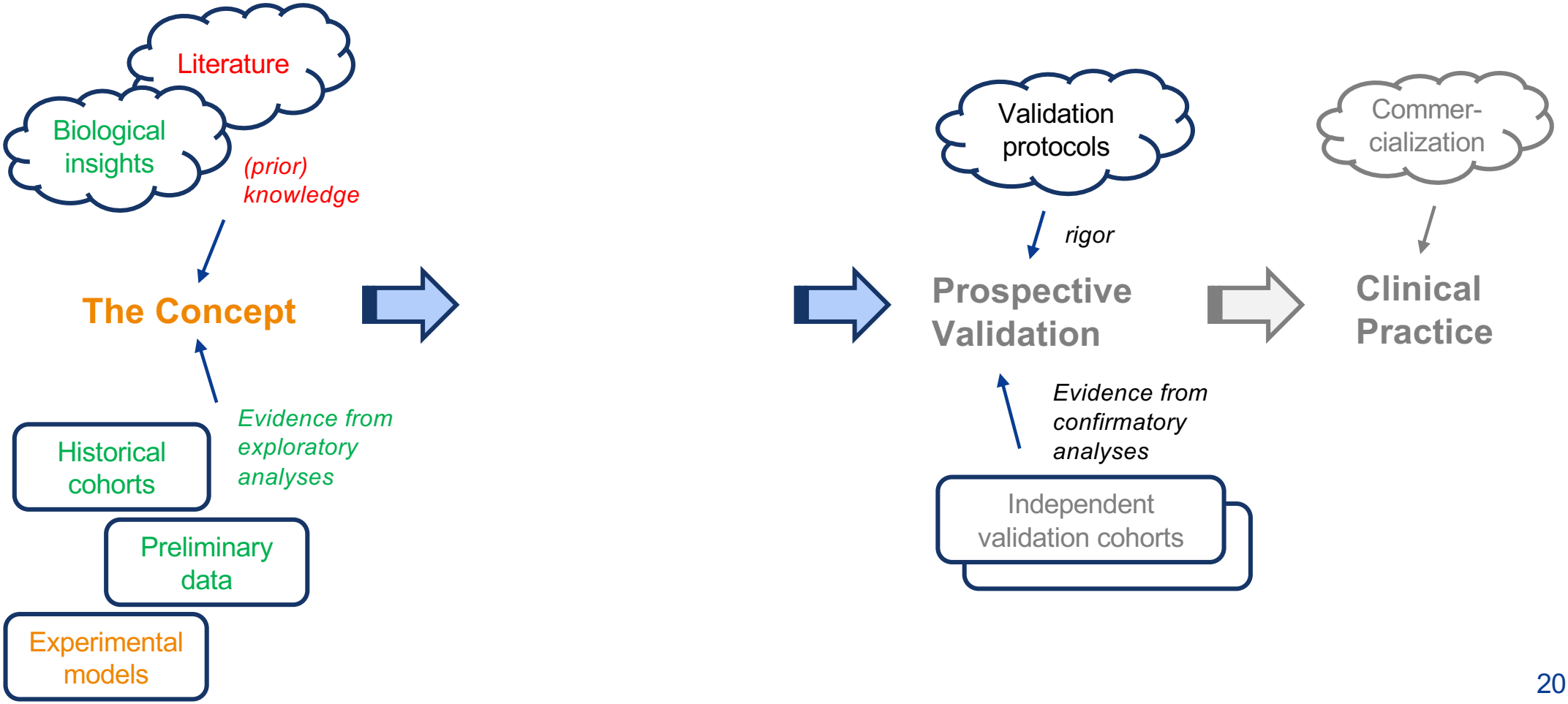
| Parameter & CutOff | Weight |
|---|---|
| LDH > 280 U/l | 0.5 |
| WBC > $4.5*10^9$ Cells/l | 0.5 |
| Age > 64 yrs | 1 |
| Cardiac comorbidity | 0.5 |
| BM Infiltration | 1 |
| Atypical cells in PB | 1 |
| Ann Arbor Stage = III or IV | 2 |
| SPD >= 3000 $mm^2$ | 2 |

**CRS-RS.5p**

# The Concept: Primarily Evidence-Driven

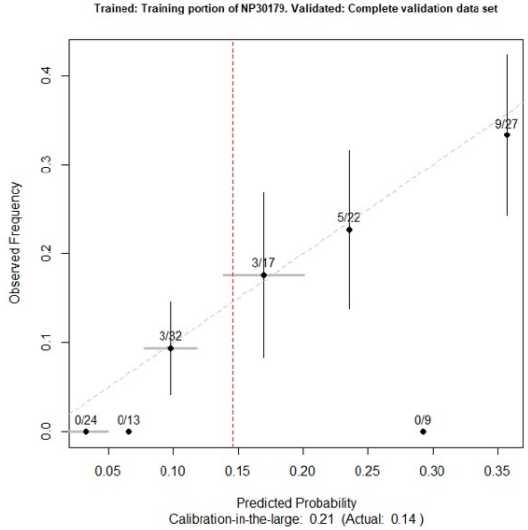


# Technical Development: -



# Prospective Trial GO43321 NIS to Validate the Model

- A prospective clinical study to determine the predictive performance of the CRS-RS.5p model
- Across clinical trials utilizing glofitamab in patients with aggressive non-Hodgkin lymphoma (excluding MCL)
- Whether the PTs are at low or high risk of developing Grade ≥ 2 CRS following the first 2.5-mg columvi dose
- N = 240 (GO43321) + 190 (2 additional prospective data collections, Roche Ph I trials)




Error bars represent:

- Y-axis: ±SE of observed frequencies and measure the precision of observed frequency summary in every bin
- X-axis: ±1.5 * standard deviation (SD) of predicted probabilities and correspond to a typical width of the distribution of predicted probabilities in every bin

The points on the plot are annotated by the number of Grade≥2 events / number of cases in the corresponding bin. Red dashed line: probability of CRS of Grade≥2 predicted by the model that corresponds to Dose=2.5mg and CRS-RS.5p=4. Gray dashed line: the diagonal that corresponds to the perfect calibration.

94% of patients from the validation cohort are CRS-RS.5p – evaluable as provide complete set of baseline risk parameter values required to estimate CRSRS.5p risk score.

- True Positive Rate (aka **Sensitivity**) **~ 0.85 .. 0.90**
- Negative Predictive Value (aka **NPV**) **~ 0.95 .. 0.98**
- Detected **fraction of low CRS risk patients ~ 50 %**

| Validation.Data.Set | N | NPV | P.NPVgreater0.9 |
|---|---|---|---|
| NP30179 Validation | 145 | 0.96 (0.024), 95%CI 0.88 to 0.99 | 0.071 |
| NP30179 Validation + GO43921 + YO42610 | 340 | 0.97 (0.013), 95%CI 0.93 to 0.99 | 0.00057 |
| IP-weighted: NP30179 Validation + GO43921 + YO42610 | 320 | 0.98 (0.011), 95%CI 0.94 to 1 | 0.00057 |
| GO43921 | 168 | 0.98 (0.017), 95%CI 0.92 to 1 | 0.0077 |
| GO43921 (prospective part) | 74 | 0.97 (0.026), 95%CI 0.86 to 1 | 0.095 |

| Validation.Data.Set | N | PPV | Sensitivity | Specificity | Prevalence | DetectionRate |
|---|---|---|---|---|---|---|
| NP30179 Validation | 145 | 0.23 (0.048) | 0.85 (0.08) | 0.54 (0.045) | 0.14 (0.029) | 0.52 (0.041) |
| NP30179 Validation + GO43921 + YO42610 | 340 | 0.18 (0.029) | 0.86 (0.056) | 0.53 (0.029) | 0.11 (0.017) | 0.51 (0.027) |
| IP-weighted: NP30179 Validation + GO43921 + YO42610 | 320 | 0.16 (0.029) | 0.9 (0.056) | 0.53 (0.029) | 0.093 (0.016) | 0.51 (0.028) |
| GO43921 | 168 | 0.17 (0.041) | 0.88 (0.083) | 0.54 (0.04) | 0.095 (0.023) | 0.5 (0.039) |
| GO43921 (prospective part) | 74 | 0.17 (0.062) | 0.86 (0.13) | 0.55 (0.061) | 0.095 (0.034) | 0.49 (0.058) |

# Model to Predict CRS after Columvi

- 5-parameter risk score was developed to predict risk of Gr2+ CRS in aNHL patients after the first dose of Columvi.
- Model uccessfully validated in a prospective, multi-centric cohort
- Utilization of the model in clinical practice requires technical development and additinal validation of the risk score

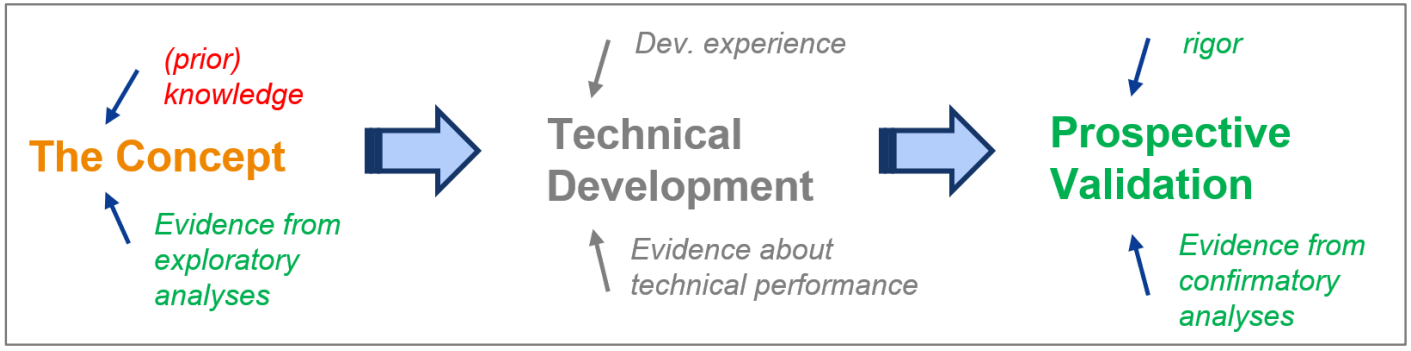
**Rather weak initial evidence:**

→ Low N

→ Training across data sets of varying exposure

**Strong performance across data sets and conditions:**

→ Yet enough room for the risk parameters to demonstrate predictive power on top of exposure



# Final Considerations

Biomarker in clinical research: **Evidence of a MoA / PD / PK may be sufficient to support research** or clinical development cycle (*biomarker data: the molecule does what it is supposed to do*) …

…**but may be insufficient when a preditive power** in a particular classification / predictiion clinical context **is desired** (*biomarker-based prediction of clinical performance*)

Why?

**1- Victory of Prior over Likelihood:**

quite weak data-driven evidence in exploratory research

*prior* *likelihood*

$$P(\theta|x) \sim P(\theta) * P(x|\theta)$$

**2- Genuine incompleteness of predictive models**

Relevant parameters may be not collected or exist as latent factors

Complex ‚playground' with several (potentially) disturbing covariates