

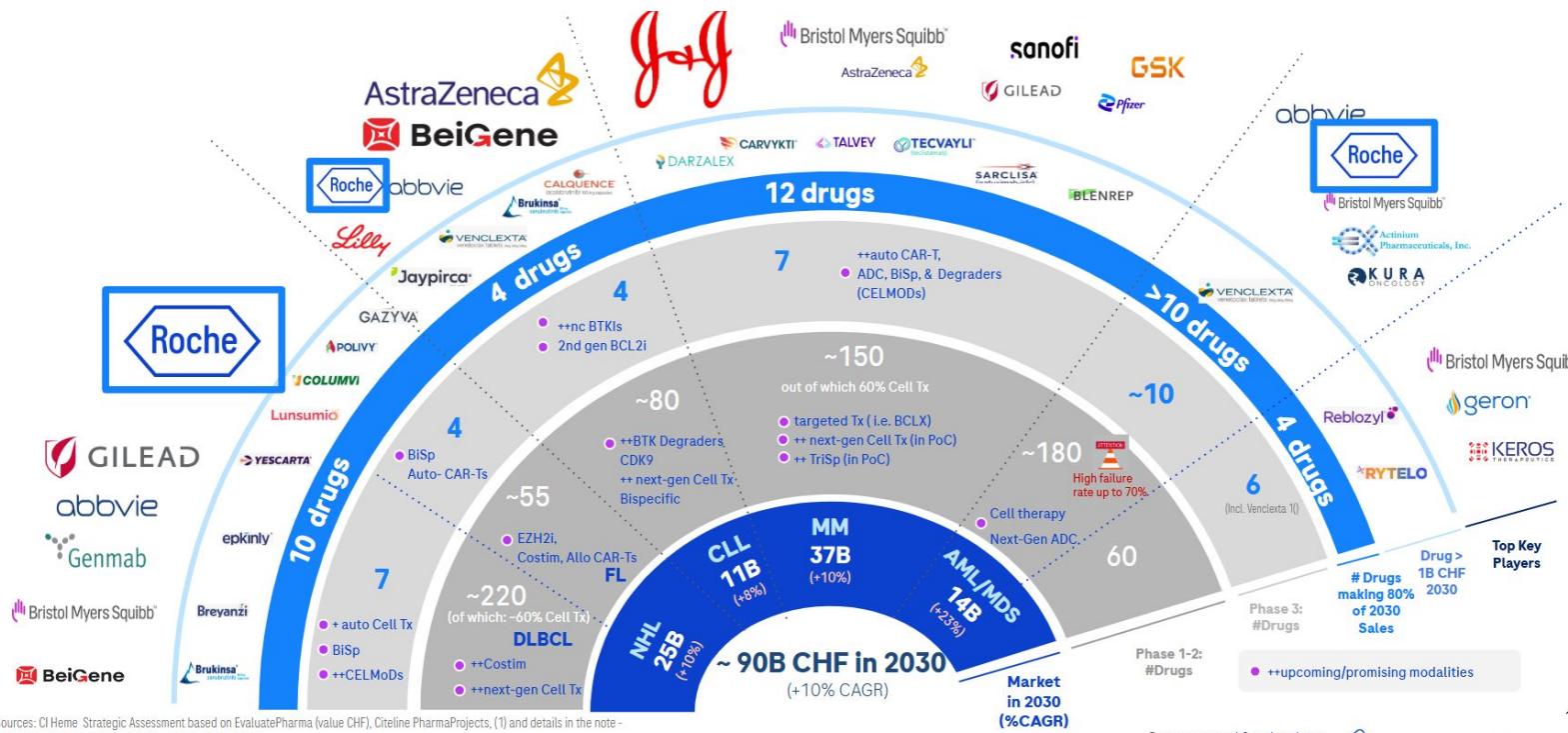
OPERATING CHARACTERISTICS OF M-PROTEIN DYNAMIC METRICS TO SUPPORT EARLY DECISIONS IN MULTIPLE MYELOMA

Pascal Chanu, Mathilde Marchand (Certara), Federico Mattiello, Mark Yan, Tom Chambers, Felipe Castro, Monica Susilo, Divya Samineni, Dale Miles, Chunze Li, Jin Jin

Clinical Pharmacology - Modeling & Simulation

September 19, 2025, Pharmacometrics in France

MULTIPLE MYELOMA (MM) REMAINS AN APPEALING MARKET

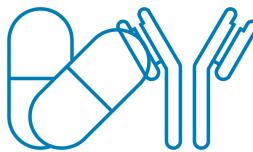


Sources: CI Heme Strategic Assessment based on EvaluatePharma (value CHF), Citeline PharmaProjects, (1) and details in the note. High level analysis, not exhaustive - no vaccine included.

Sources used for the data

Oct 2024

EARLY DEVELOPMENT SITUATION IN MULTIPLE MYELOMA (MM)



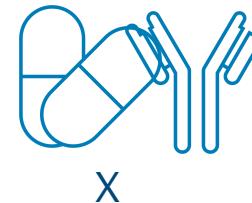
Combination therapy **X** in multiple myeloma
Phase 1b data, N=30,
2 dose levels, no control,
endpoint:
ORR i.e. VGPR rate

External data from YODA



RWD flatiron

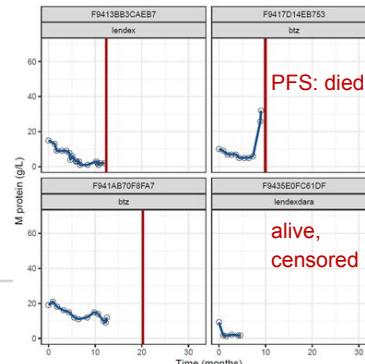
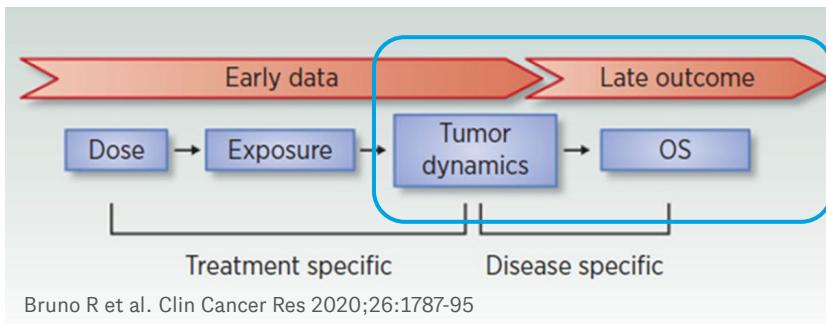
MM enabling decision framework



Phase 3 study:
X
versus
Standard of Care:
daratumumab+polatuzumab+
dexamethasone (DPd),
endpoint: PFS

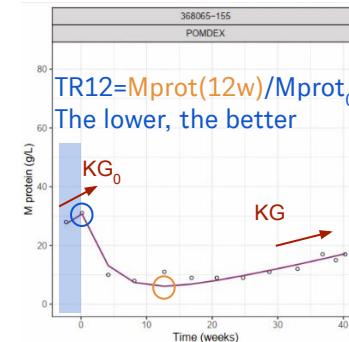
MODELING & SIMULATION FRAMEWORK IN MULTIPLE MYELOMA

Model-based tumor dynamics metrics are biomarkers capturing treatment effect



KG, TR12...
linked to PFS

Application to Multiple Myeloma

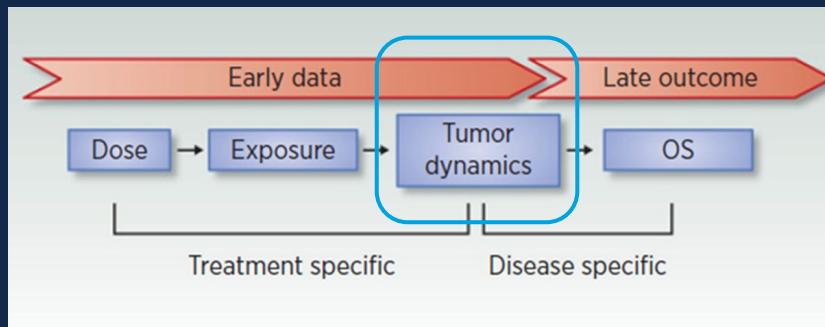


$$t < 0 \quad Mprot(t) = Mprot_0 \cdot e^{KG_0 \cdot t}$$

$$t \geq 0 \quad Mprot(t) = Mprot_0 \cdot (e^{KG \cdot t} + e^{-KS \cdot t} - 1)$$

Stein et al. Clin Cancer Res 2011;17:907-17

OPERATING CHARACTERISTICS OF M-PROTEIN DYNAMIC METRICS



OPERATING CHARACTERISTICS OF TGI METRICS

Tumor Dynamic Model-Based Decision Support for Phase Ib/II Combination Studies: A Retrospective Assessment Based on Resampling of the Phase III Study IMpower150

René Bruno¹, Mathilde Marchand², Kenta Yoshida³, Phyllis Chan³, Haocheng Li⁴, Wei Zou⁵, Francois Mercier⁶, Pascal Chanu⁷, Benjamin Wu³, Anthony Lee⁵, Chunze Li³, Jin Y. Jin³, Michael L. Maitland^{8,9}, Martin Reck¹⁰, and Mark A. Socinski¹¹



ABSTRACT

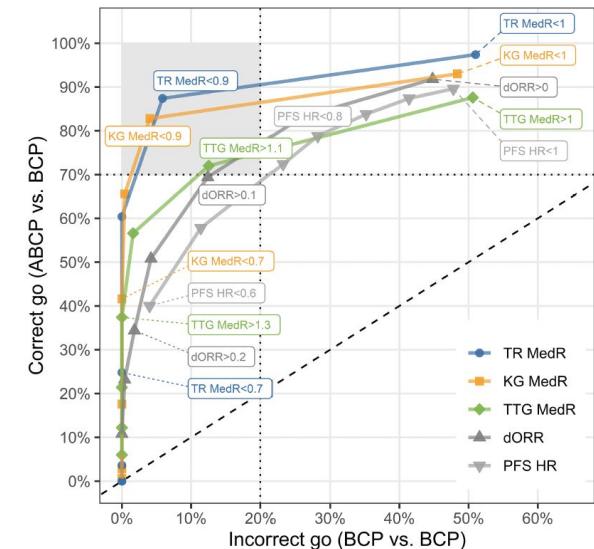
Purpose: Model-based tumor growth inhibition (TGI) metrics are increasingly incorporated into go/no-go decisions in early clinical studies. To apply this methodology to new investigational combinations requires independent evaluation of TGI metrics in recently completed Phase III trials of effective immunotherapy.

Patients and Methods: Data were extracted from IMpower150, a positive, randomized, Phase III study of first-line therapy in 1,202 patients with non-small cell lung cancer. We resampled baseline characteristics and longitudinal sum of longest diameters of tumor lesions of patients from both arms, atezolizumab+bevacizumab+chemotherapy (ABCP) versus BCP, to mimic Phase Ib/II studies of 15 to 40 patients/arm with 6 to 24 weeks follow-up. TGI metrics were estimated using a bi-exponential TGI model. Effect sizes were calculated as TGI metrics ratio of medians (MedR), objective response rate (ORR) difference (d),

and progression-free survival (PFS), hazard ratio (HR) between arms. Correct and incorrect go decisions were evaluated as the probability to achieve desired effect sizes in ABCP versus BCP and BCP versus BCP, respectively, across 500 replicated subsamples for each design.

Results: For 40 patients/24 weeks follow-up, correct go decisions based on probability tumor growth rate (KG) MedR <0.90, dORR >0.10, and PFS HR <0.70 were 83%, 69%, and 58% with incorrect go decision rates of 4%, 12%, and 11%, respectively. For other designs, the ranking did not change with TGI metrics consistently overperforming RECIST endpoints. The predicted overall survival (OS) HR was around 0.80 in most of the scenarios investigated.

Conclusions: Model-based estimate of KG MedR is an exploratory endpoint that informs early clinical decisions for combination studies.

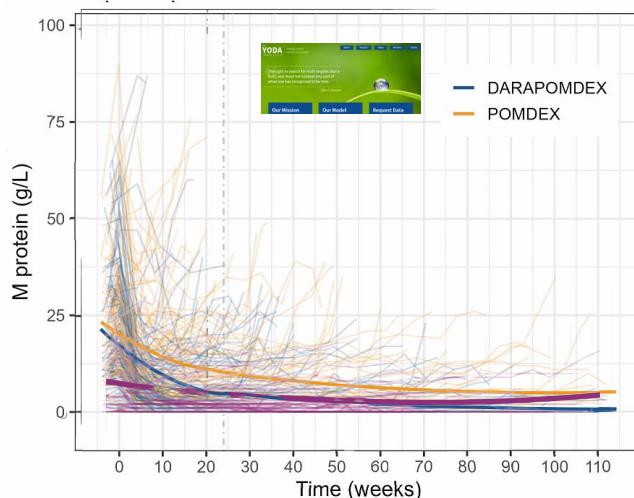


APOLLO PATIENT POPULATION IS REPRESENTATIVE

APOLLO: model development, DPd
versus Pd
Dimopoulos et al. Lancet Oncol 2021;22:801-12

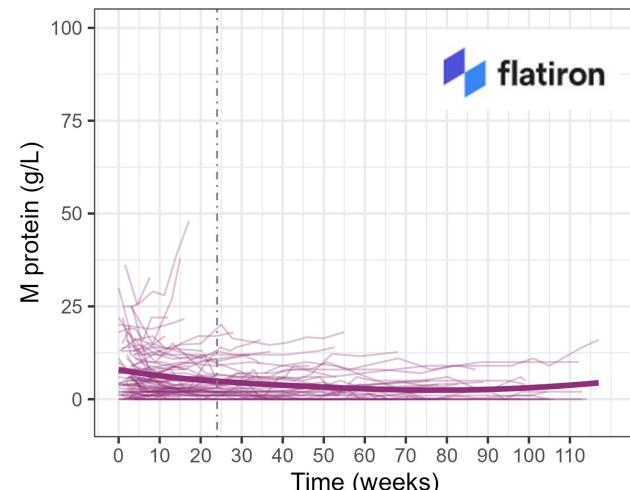
Daratumumab plus pomalidomide and dexamethasone
versus pomalidomide and dexamethasone alone in
previously treated multiple myeloma (APOLLO):
an open-label, randomised, phase 3 trial

Meletios A Dimopoulos, Evangelos Terpos, Mario Boccadoro, Sosana Delimpasi, Meral Beksoz, Eirini Karodimitou, Philippe Moreau, Luca Baldini, Argiro Symeonidis, Jelena Bilo, Albert Orrix, Maria-Victoria Mateos, Hernandez-Insiste, Ioannis Orfanidis, Tahamtan Ahmadi, Joe Uktropic, Tobias Kornifchenko, Jordan M Schechter, Yaping Qiu, Hima Amin, Jessica Vermeulen, Robin Carson, Pieter Sonneveld, for the APOLLO Trial Investigators*



DPd in
APOLLO
considered as
a relevant
control for X
assessment

RWD of the population of interest shows a
good match with DPd arm from APOLLO



RWD from Flatiron Health, an oncology-based electronic
health record (EHR)-derived de-identified database

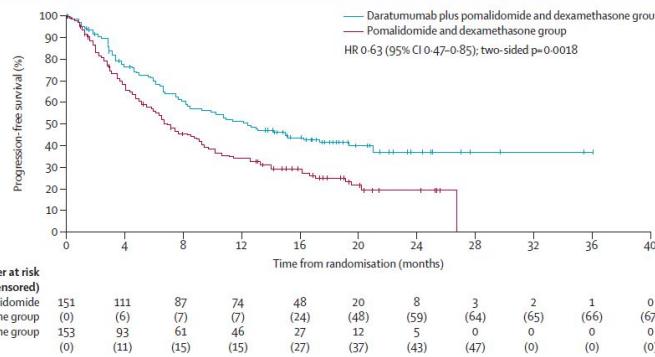
Genentech
A Member of the Roche Group

DATA: APOLLO & CASTOR CLINICAL TRIALS

APOLLO: model development, DPd versus Pd
Dimopoulos et al. Lancet Oncol 2021;22:801-12

Daratumumab plus pomalidomide and dexamethasone versus pomalidomide and dexamethasone alone in previously treated multiple myeloma (APOLLO): an open-label, randomised, phase 3 trial

Meletios A Dimopoulos, Evangelos Terpos, Mario Boccadoro, Sosana Dellimpasi, Meral Beksaç, Eirini Katsodrati, Philippe Moreau, Luca Baldini, Argiris Syrigonis, Jefena Bila, Albert Orfao, Maria-Victoria Mateos, Hermann Einsle, Ioannis Orfanidis, Taharman Ahmadi, Jon Ukkopen, Tobias Kampfner, Jordan M Schechter, Yanping Qiu, Hima Amin, Jessica Vermaelen, Robin Carson, Pieter Sonneveld, for the APOLLO Trial Investigators*



DPd in APOLLO considered as a relevant control for X assessment

For assessing operating characteristics of M-protein dynamics: DPd will be the experimental arm and Pd the control

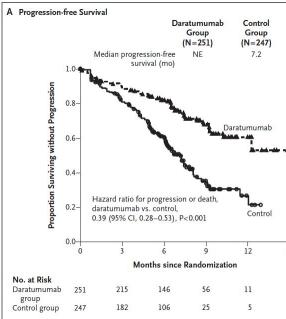
CASTOR: validation, DVd versus Vd
Palumbo et al. N Engl J Med 2016;375:754-66

THE NEW ENGLAND JOURNAL OF MEDICINE



Daratumumab, Bortezomib, and Dexamethasone for Multiple Myeloma

Antonio Palumbo, M.D., Asher Chanian-Khan, M.D., Katja Weisel, M.D., Ajay K. Nooka, M.D., Tamas Masszi, M.D., Meral Beksaç, M.D., Ivan Spicka, M.D., Vania Hungria, M.D., Markus Munder, M.D., Maria V. Mateos, M.D., Tomer M. Mark, M.D., Ming Qi, M.D., Jordan Schechter, M.D., Hima Amin, B.S., Xiang Qin, M.S., William Deraadt, Ph.D., Taharman Ahmadi, M.D., Andrew Spencer, M.D., and Pieter Sonneveld, M.D., for the CASTOR Investigators*



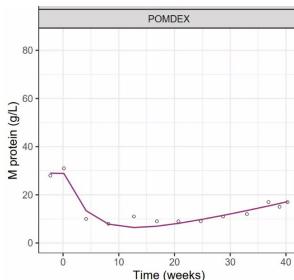
MODEL SELECTION TO ASSESS OPERATING CHARACTERISTICS (OC) ON APOLLO DATA

Model development performed in **nlmixr²**

$$t < 0 \quad Mprot(t) = Mprot_0 \cdot e^{KG \cdot t}$$

$$t \geq 0 \quad Mprot(t) = Mprot_0 \cdot (e^{KG \cdot t} + e^{-KS \cdot t} - 1)$$

Parameter	Estimate	SE	RSE (%)	Back-transformed	Unit	Shrinkage (%)
Fixed Effects						
ln(KG_DPD)	-7.50	0.0107	0.1	0.000553	day ⁻¹	
ln(KG_Pd)	-6.49	0.0118	0.2	0.00152	day ⁻¹	
ln(KS_DPD)	-3.78	0.0114	0.3	0.0228	day ⁻¹	
ln(KS_Pd)	-4.33	0.0107	0.2	0.0132	day ⁻¹	
TS0	2.80	0.00544	0.2	16.4	g/L	
Random Effects						
KG	82			%CV	66.8	
KS	81			%CV	6.1	
TS0	81			%CV	0.6	
Residual variability						
Res add.	2.78					
Run: noKG objective function value: 13070.03						

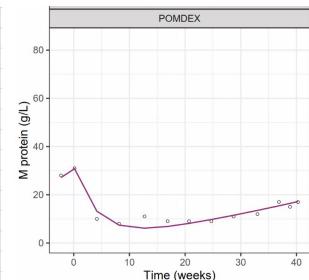


Estimating a different KG₀ before treatment improves the model

$$t < 0 \quad Mprot(t) = Mprot_0 \cdot e^{KG_0 \cdot t}$$

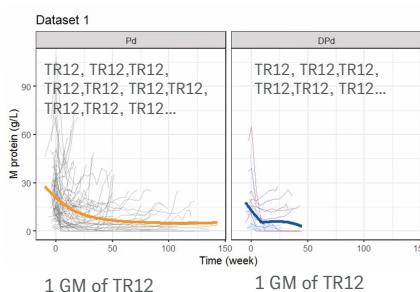
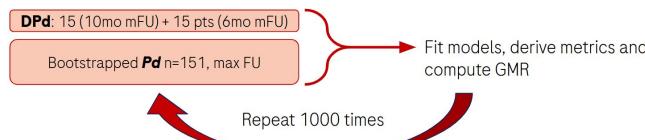
$$t \geq 0 \quad Mprot(t) = Mprot_0 \cdot (e^{KG \cdot t} + e^{-KS \cdot t} - 1)$$

Parameter	Estimate	SE	RSE (%)	Back-transformed	Unit	Shrinkage (%)
Fixed Effects						
ln(KG ₀)	-6.79	0.157	2	0.00112	day ⁻¹	
ln(KG_DPD)	-7.02	0.0808	1	0.000894	day ⁻¹	
ln(KG_Pd)	-6.36	0.0852	1	0.00173	day ⁻¹	
ln(KS_DPD)	-4.23	0.123	3	0.0146	day ⁻¹	
ln(KS_Pd)	-4.36	0.127	3	0.0128	day ⁻¹	
TS0	2.82	0.0893	3	16.8	g/L	
Random Effects						
KG ₀	165			%CV	48.5	
KG	82			%CV	61.3	
KS	80			%CV	6.4	
TS0	81			%CV	3.9	
Residual variability						
Res add.	2.93					
Run: KG0llV objective function value: 13041.53						



ASSESSMENT OF OPERATING CHARACTERISTICS (OC) APRIL 2025

True GO: DPd vs. Pd

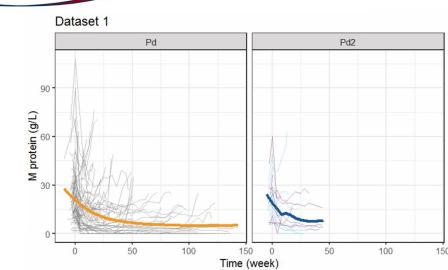
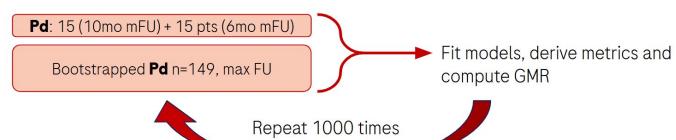


1 GM of TR12

1 GMR exp/control
The lower, the better

% of achieving a certain target GMR
e.g. <0.75

False GO: Pd vs. Pg



- KG has poor OCs, **what's wrong?**
- TR8, TR12 are good metrics (better than TR6 and TR24)
- $p(\text{TR GMR} < 0.75)$ or $p(\text{TR GMR} < 0.80)$ as gating criteria

IMPORTANCE OF MODEL'S ROBUSTNESS WHEN ASSESSING OC

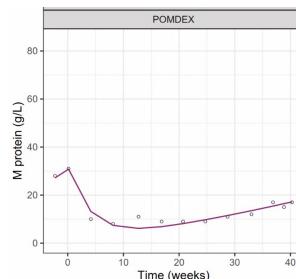
$$t < 0 \quad M_{prot}(t) = M_{prot_0} \cdot e^{KG_0 \cdot t}$$

$$t \geq 0 \quad M_{prot}(t) = M_{prot_0} \cdot (e^{KG \cdot t} + e^{-KS \cdot t} - 1)$$

Estimating a different KG0 before treatment improves the model

Parameter	Estimate	SE	RSE (%)	Back-transformed	Unit	Shrinkage (%)
Fixed Effects						
In(KG0)	-6.79	0.157	2	0.00112	day ⁻¹	
In(KG_DPD)	-7.02	0.0808	1	0.000894	day ⁻¹	
In(KG_Pd)	-6.36	0.0852	1	0.00173	day ⁻¹	
In(KS_DPD)	-4.23	0.123	3	0.0146	day ⁻¹	
In(KS_Pd)	-4.36	0.127	3	0.0128	day ⁻¹	
TS0	2.82	0.0893	3	16.8	g/L	
Random Effects						
KG0	165				%CV	48.5
KG	82				%CV	61.3
KS	80				%CV	6.4
TS0	81				%CV	3.9
Residual variability						
Res.add.	2.93				g/L	

Run: KG0IIV objective function value: 13041.53



On small subsamples, the typical value of KG0 is identifiable but likely not its random effect

Parameter	Estimate	SE	RSE (%)	Back-transformed	Unit	Shrinkage (%)
Fixed Effects						
In(KG0)	-5.96	0.127	2	0.00258	day ⁻¹	
In(KG_DPD)	-7.63	0.0792	1	0.000486	day ⁻¹	
In(KG_Pd)	-6.37	0.0966	2	0.00171	day ⁻¹	
In(KS_DPD)	-3.80	0.117	3	0.0224	day ⁻¹	
In(KS_Pd)	-4.32	0.129	3	0.0133	day ⁻¹	
TS0	2.68	0.0769	3	14.6	g/L	
Random Effects						
KG0	NA				%CV	NA
KG	82				%CV	60.3
KS	76				%CV	5.3
TS0	70				%CV	7.9
Residual variability						
Res.add.	3.27				g/L	

Run: KG0nollV objective function value: 13326.31

This model makes more sense:

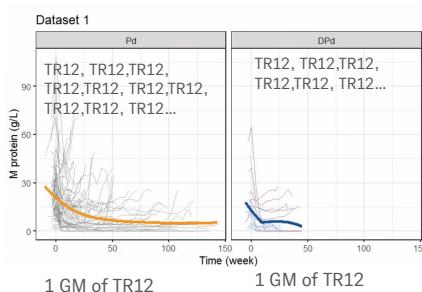
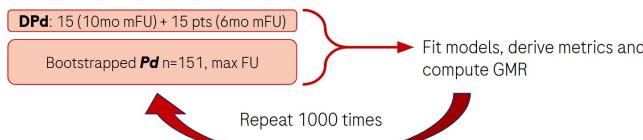
KG0 > KG_Pd > KG_DPD

[0.00201;0.00331] [0.00142;0.00207] [0.000416 ; 0.000567]

Let's redo OCs with this model

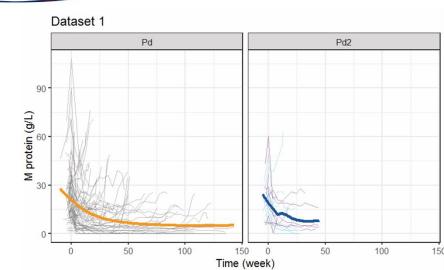
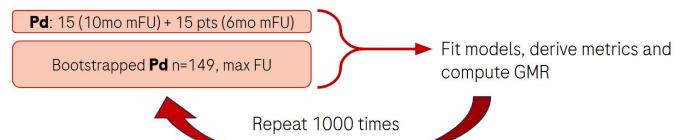
RE-ASSESSMENT OF OPERATING CHARACTERISTICS JULY 2025

True GO: DPd vs. Pd



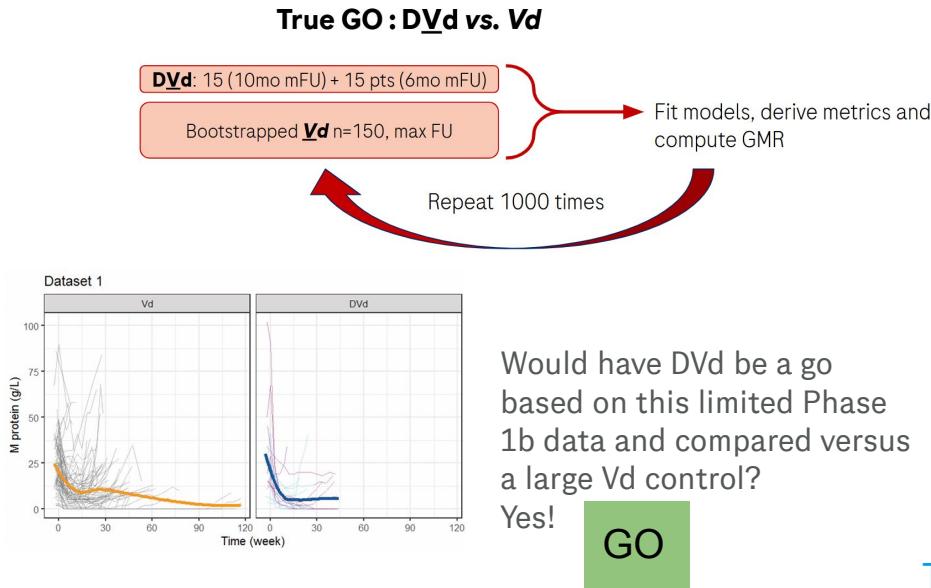
GM of TR12 1 GM of TR12
↓ ↓
1 GMR exp/control
The lower, the better
X 1000
% of achieving a
certain target GMR
e.g. <0.75

False GO: Pd vs. Pd



- Results make sense with a less parameterized model
- KG is now a decent metric
- TR8, TR12 are good metrics
- $p(\text{TR GMR} < 0.70)$ as gating criteria

LET'S TEST THE APPROACH ON THE CASTOR TRIAL



Probability of achieving the gating criteria

Experimental (DVd) versus Control (Vd): True GO

Metric	p(GMR<0.70)%	p(GMR<0.75)%
KG	94.1	96.9
TR8	74.3	77.4
TR12	87.5	90.2

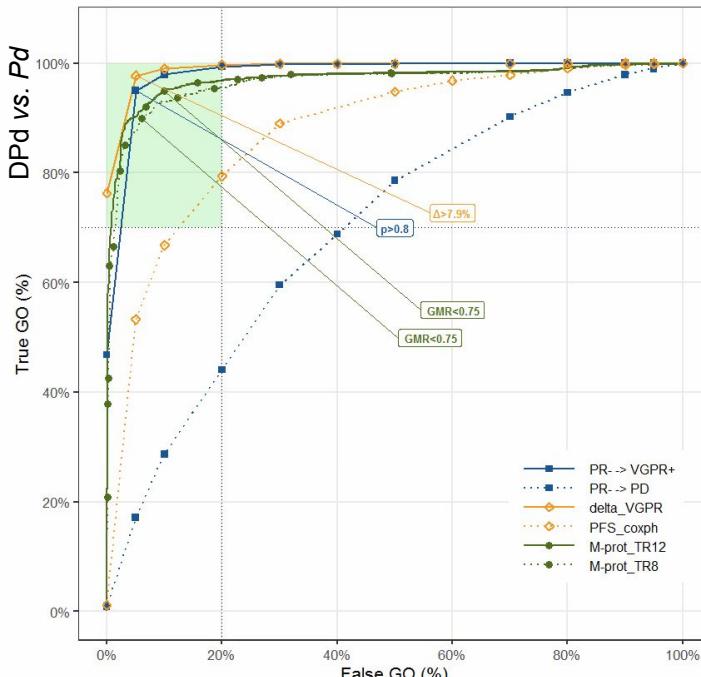
Control (Vd) versus Control (Vd)/ False GO

Metric	p(GMR<0.70)%	p(GMR<0.75)%
KG	11.8	16.3
TR8	3.8	8.2
TR12	7.6	12.1

The metric and the gating criteria are validated
We can use this approach to ungate a Phase 3 with treatment X in comparison to DPd

COMPARISON M-PROTEIN DYNAMICS VERSUS OTHER APPROACHES

Operating Characteristics (OCs) for the various metrics based on APOLLO



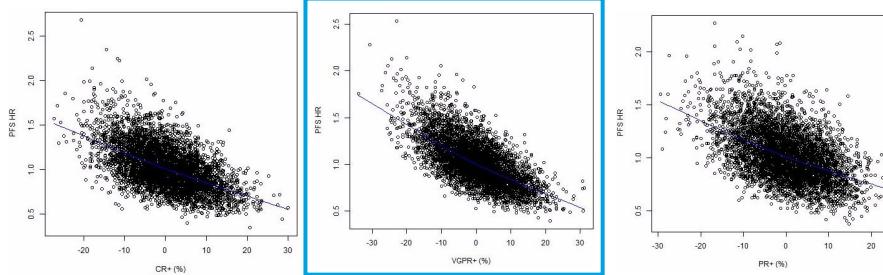
Method	Data	Assessment	Key assets
-VGPR+ (Bayesian approach) - Δ VGPR+	Observed	X	VGPR: Primary analysis - Association between Δ VGPR+ and PFS HR
Multistate model	Model-predicted	X vs control (DPd from APOLLO)	-Leverages longitudinal ORR -Transition PR \rightarrow VGPR has good OCs
M-protein dynamic model	Model-predicted	X (M-protein evaluable) vs control (DPd from APOLLO)	-Leverages longitudinal continuous M-protein data -M-protein ratios to baseline at weeks 8 & 12 have good OCs

ΔVGPR+ and the 2 model-based approaches demonstrate very good OCs while PFS has relatively poor OCs in early development settings (few subjects, short follow-up)

Courtesy: Mark Yan, Federico Mattiello and MM decision enabling framework
Multistate model and Bayesian analysis

LET'S PRESSURE TEST, WHAT IS THE ASSOCIATION WITH PFS?

- Meta-analysis based on 21 randomized Ph2/3 trials published after 2014
 - VGPR+ rate has the strongest association (among other response-based endpoints) with median PFS
 - No strong association found between delta in response rate and PFS HR probably due to heterogeneity in patient populations
- Patient-level data from APOLLO (DPd arm, N = 146)
 - Low to moderate correlation between response and PFS; VGPR+ is superior to PR+ and CR+
 - M-protein dynamic metric logKG shows similar association than delta in VGPR+

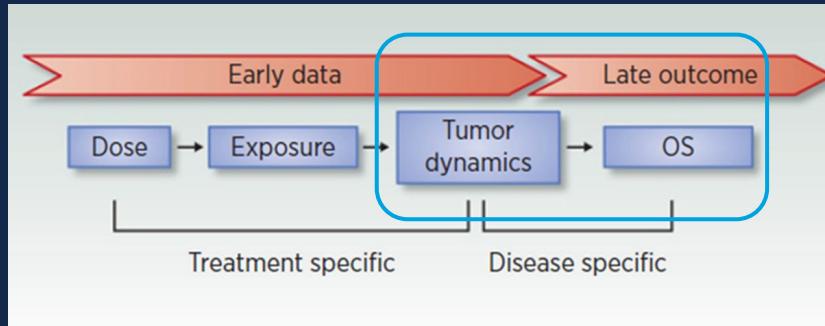


KG: growth constant,
TR12: Tumor ratio to
baseline at week 12,
TTG: Time To Growth...

	APOLLO (DPd)	
Endpoint	Correlation coef.	R-sq
Delta in VGPR+	-0.68	0.46
logKG	0.66	0.44
KG	0.64	0.41
TR24	0.59	0.34
TR18	0.52	0.27
TTG	-0.49	0.24
TR12	0.40	0.16
TR8	0.31	0.10

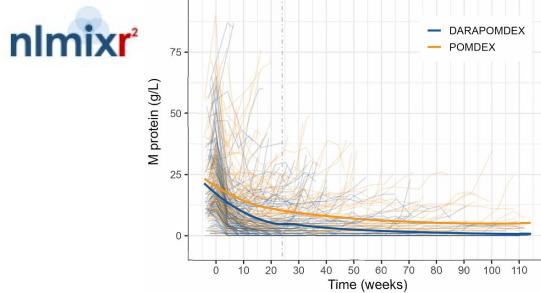
Courtesy: Mark Yan and MM decision enabling framework

PROBABILITY OF TECHNICAL SUCCESS OF PHASE 3 BASED ON PFS SIMULATIONS



TGI-PFS MODEL DEVELOPMENT ON APPolo

1. Derive M-protein dynamic metrics



$$t < 0 \quad Mprot(t) = Mprot_0 \cdot e^{KG_0 \cdot t}$$

$$t \geq 0 \quad Mprot(t) = Mprot_0 \cdot (e^{KG \cdot t} + e^{-KS \cdot t} - 1)$$

Parameter	Estimate	SE	RSE (%)	Back-transformed	Unit	Shrinkage (%)
Fixed Effects						
ln(KG0)	-5.96	0.127	2	0.00258	day ⁻¹	
ln(KG_DPd)	-7.63	0.0792	1	0.000486	day ⁻¹	
ln(KG_Pd)	-6.37	0.0966	2	0.00171	day ⁻¹	
ln(KS_DPd)	-3.80	0.117	3	0.0224	day ⁻¹	
ln(KS_Pd)	-4.32	0.129	3	0.0133	day ⁻¹	
TS0	2.68	0.0769	3	14.6	g/L	
Random Effects						
KG0	NA			%CV	NA	
KG	82			%CV	60.3	
KS	76			%CV	5.3	
TS0	70			%CV	7.9	
Residual variability						
Res.add.	3.27				g/L	
Run: KG0nolV objective function value: 13326.31						

2. Univariate survival analysis

	Score	p.LRT	N	Sign
logKG	230.7	0.0000	239	+
TTG	89.7	0.0000	239	-
KG	77.2	0.0000	239	+
TR6	31.6	0.0000	239	+
pred_BSLD	25.3	0.0000	239	+
TR8	24.5	0.0000	239	+
ALBU	17.1	0.0000	239	-
HGB	15.3	0.0001	239	-
TR12	13.4	0.0003	239	+
iss	11.6	0.0006	239	+
iss1	10.9	0.0010	239	-
Obs_BSLD	10.9	0.0010	239	+
TR18	8.0	0.0047	239	+
cytrisk	7.2	0.0271	175	-
TR24	6.8	0.0090	239	+
iss3	5.9	0.0148	239	+
SECR	5.2	0.0230	239	+
logKS	3.6	0.0590	239	-
rengrpnn	2.8	0.0948	239	-
LDH	2.4	0.1235	239	+
iss2	2.1	0.1502	239	+
BIL	1.7	0.1882	239	+
ALP	1.4	0.2330	239	+
CRCL	1.3	0.2577	239	-
TPRO	1.2	0.2670	239	+
KS	0.9	0.3479	239	-
ALT	0.7	0.4090	239	+
age	0.5	0.4940	239	-
BECOG0123	0.3	0.5928	239	+
BWT	0.2	0.6596	239	+
AST	0.1	0.7695	239	-
SEX	0.1	0.7386	239	-
racegrln	0.1	0.7089	239	+

M-protein dynamic metrics

Tested in multivariate survival analysis

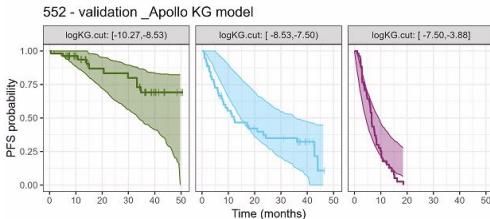
3. Multivariate parametric survival model

Term	Estimate	std.error	Z-value	p-value	Lower CI	Upper CI
(Intercept)	-5.210	0.563	-9.255	<2e-16	-6.313	-4.106
logKG	-0.770	0.051	-15.226	<2e-16	-0.869	-0.671
HGB	0.193	0.041	4.742	0	0.113	0.273
Log(scale)	-0.110	0.055	-2.010	0.0444		

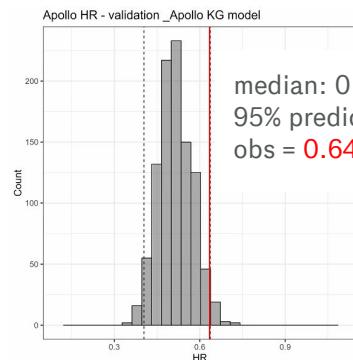
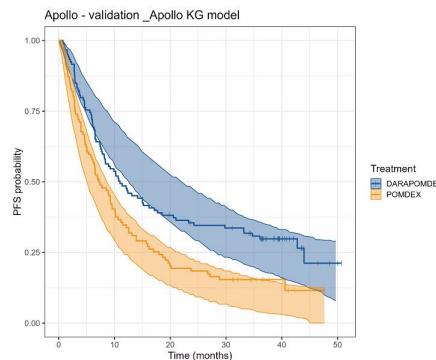
PFS is longer with slower growth rate and higher baseline hemoglobin (HGB)

TGI-PFS MODEL VALIDATION ON APOLLO & CASTOR

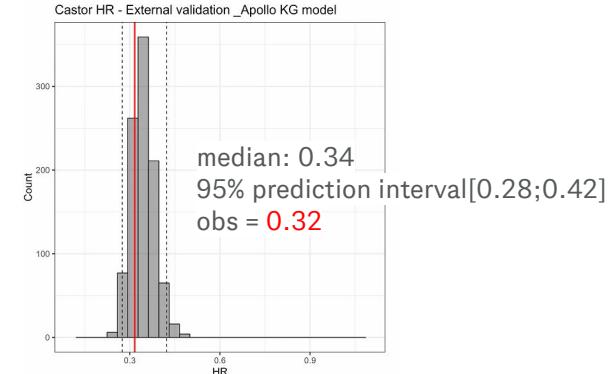
PFS model qualification on APOLLO



The model predicts quite well PFS in the different quartiles of log(KG)



External Validation on CASTOR



PFS model validated, it can be used to simulate Phase 3 trial with treatment X and assess its probability of technical success (PTS)

PROBABILITY OF TECHNICAL SUCCESS OF PHASE 3



Phase 3 study:

X

versus

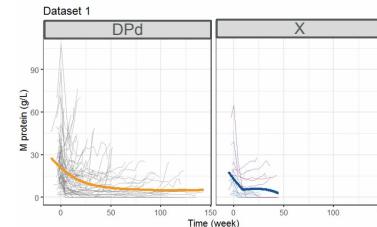
Standard of Care:

daratumumab+polamidomide+
dexamethasone (DPd),

endpoint: PFS

X: 15 (10mo mFU) + 15 pts (6mo mFU)

DPd: complete arm n=120, max FU



PFS simulations of Phase 3 design

160 vs. 160 patients

1000 trials

1000 HR, 1000 Logrank tests

- 1 median HR and 95% prediction interval
- 1 probability of Logrank tests with $p < 0.05$
- 1 probability of $HR < \text{target}$

RISK ASSOCIATED WITH DECISION MAKING BASED ON LIMITED DATA

DPd: 15 (10mo mFU) + 15 pts (6mo mFU)

Pd: complete arm n=120, max FU

Distribution of median HR:
median 0.69 [0.25;1.09]
Observed HR: 0.63

Distribution of probability of success: median 71% [6;100]

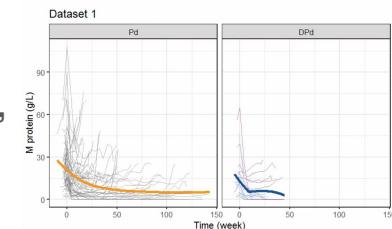
PTS assessment highly depends on the small experimental arm data

In early development, we only have N=1 set of Phase 1b data to make the decision

X500

each with a different DPd arm

Fit M-protein data,
get log(KG),
add HGB



PFS simulations of APOLLO design
150 vs. 150 patients

200 trials

200 HR, 200 Logrank tests

- 1 median HR
- 1 probability of Logrank tests with $p < 0.05$

RISK ASSOCIATED WITH DECISION MAKING BASED ON LIMITED DATA

DVd: 15 (10mo mFU) + 15 pts (6mo mFU)

Vd: complete arm n=250, max FU

Distribution of median HR:
median 0.58 [0.20;0.91]
Observed HR: 0.39

Distribution of probability of success:
median 99.5% [19.0;100.0]

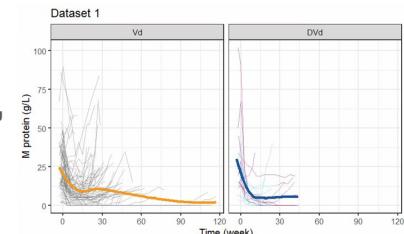
PTS assessment highly depends on the small experimental arm data

In early development, we only have N=1 set of Phase 1b data to make the decision

X500

each with a different DPd arm

Fit M-protein data,
get log(KG),
add HGB



PFS simulations of CASTOR design
250 vs. 250 patients

200 trials

200 HR, 200 Logrank tests

- 1 median HR
- 1 probability of Logrank tests with $p < 0.05$

CONCLUSIONS

- Assessing Operating Characteristics is a valuable approach to identify what is the metric, the target and the gating criteria to rationalize early decision making in a quantitative manner
 - Importance of the robustness of the model used to analyze the 1000 subsamples
 - The metric selected should ideally be well associated to the primary endpoint used in the Phase 3 study we want to ungate
- The traditional TGI-PFS (OS) approach can also be used to simulate Phase 3 outcomes and assess its probability of technical success (PTS)
 - As expected, there is a very high weight on the small Phase 1b data on the decision, i.e. the variability of the PTS is huge which leads to risky decision, we should consider larger Phase 1b trials to de-risk Phase 3 Go/No Go decisions

THANK YOU

Genentech

A Member of the Roche Group