Translational Modeling of Relapsing Mouse Models to Inform Regimen Selection in Tuberculosis

James Clary, PharmD

Assistant Director, Pharmacometrics

Please note: this presentation, including questions from the audience, is being recorded and may be made available.

Tuberculosis as global health challenge

Tuberculosis (TB) continues to remain a major global health challenge

- Estimated 10 million cases in 2019
- Despite emerging treatments and increased investment, target milestones towards decreasing the global burden of TB are still not being achieved
  - COVID-19 has set us back further, contributing to an increase in TB-related morbidity and mortality in 2020

Global trends in the estimated number of TB deaths (left) and the mortality rate (right), 2000–2020

Shaded areas represent uncertainty intervals. The horizontal dashed line shows the 2020 milestone of the End TB Strategy.

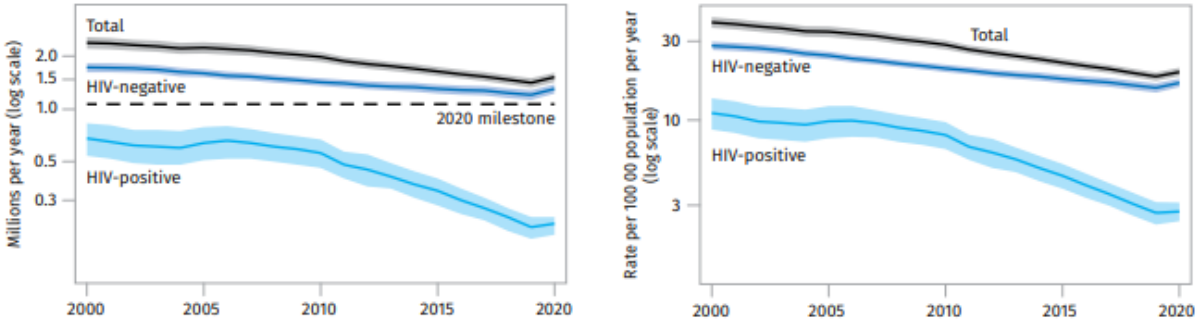


Figure from: Global tuberculosis report 2021. Geneva: World Health Organization; 2021. Licence: CC BY-NC-SA 3.0 IGO.

Investment in TB-Related R&D Efforts

The Bill & Melinda Gates Foundation is a major supporter of efforts to improve treatments and outcomes for patients with TB

https://www.gatesfoundation.org/our-work/programs/global-health/tuberculosis

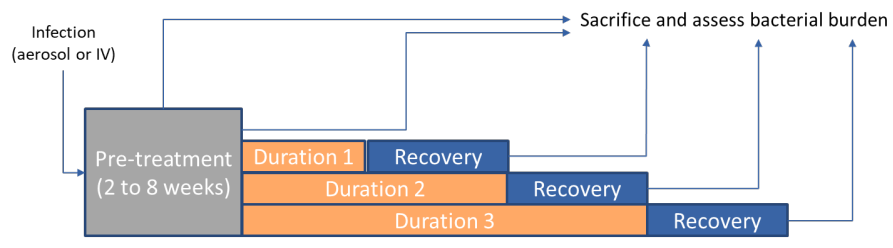
BILL & MELINDA GATES foundation

- Areas of Focus
  - More effective drug regimens
  - New diagnostic tools
  - Improved vaccines

SimulationsPlus
Cognigen | DILIsym Services | Lixoft

Relapsing Mouse Model Studies

- **The curative potential of novel anti-TB drugs and regimens is often assessed via "relapsing mouse model" (RMM) studies in BALB/c mice**



- **Analysis of RMM studies relies on simple statistical calculations comparing proportions of mice exhibiting relapse following treatment with selected regimens at limited treatment durations**

<u>Why is this important</u>?
→ RMM studies are highly influential in regimen prioritization for further study and often inform regimen selection for clinical evaluation

Treatment	Proportion of mice relapsing after treatment for:			
	1 month	1.5 months	2 months	2.5 months
Regimen A	--	--	7/15 (47%)	2/15 (47%)
Regimen B	--	14/15 (93%)	10/15 (47%)	8/15 (53%)
Regimen C	--	5/15 (33%)	0/15 (0%)	--
Regimen D	--	8/15 (53%)	6/15 (40%)	--
Regimen E	4/15 (27%)	0/15 (0%)	--	--

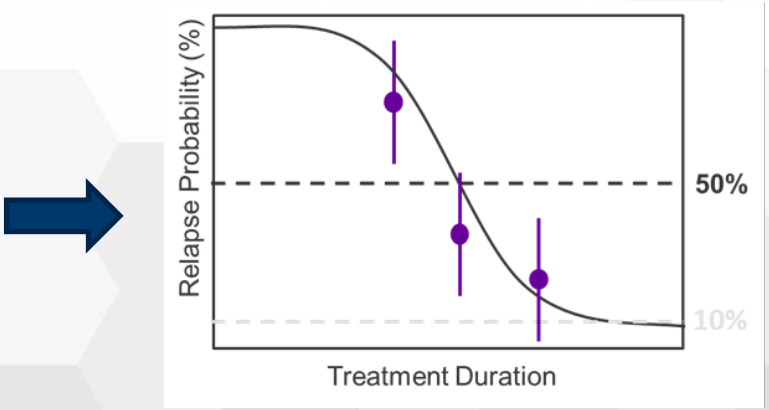
RMM Studies → We can do better!

The focus on assessing relative regimen efficacy based on proportions of relapsing mice at limited treatment durations *does not* provide answers to key questions such as:

"How long do we need to treat with Regimen A to achieve a high probability of cure?"

"How does the time to X% relapse probability compare for novel regimens vs. established regimens?"

Treatment	Proportion of mice relapsing after treatment for:			
	1 month	1.5 months	2 months	2.5 months
Regimen A	--	--	7/15 (47%)	2/15 (13%)
Regimen B	--	14/15 (93%)	10/15 (47%)	8/15 (53%)
Regimen C	--	5/15 (33%)	0/15 (0%)	--
Regimen D	--	8/15 (53%)	6/15 (40%)	--
Regimen E	4/15 (27%)	0/15 (0%)	--	--



Simulations to Inform RMM Study Design

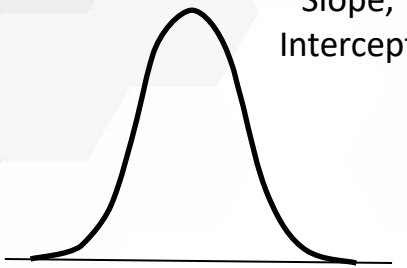
As part of BMGF's TB-related collaborations, a large series of RMM studies were being proposed by a partner organization

- Given the significant investment in the proposed RMM experiments, a thorough assessment of the proposed study design was proposed
- A previously developed model* of RMM efficacy was considered a valuable tool to evaluate study designs through simulations
- The objectives were two-fold:
  - Evaluate proposed design for variety of hypothetical regimens to quantify the design's ability to determine key metrics (Time to 50% relapse [T_{50}] and Time to 5% relapse [T5]) for regimen comparison
  - Recommend alternative study designs that maximize informativeness and decrease overall number of mice utilized

*Berg et al 2022 Jan 31;AAC0179321. doi: 10.1128/AAC.01793-21.

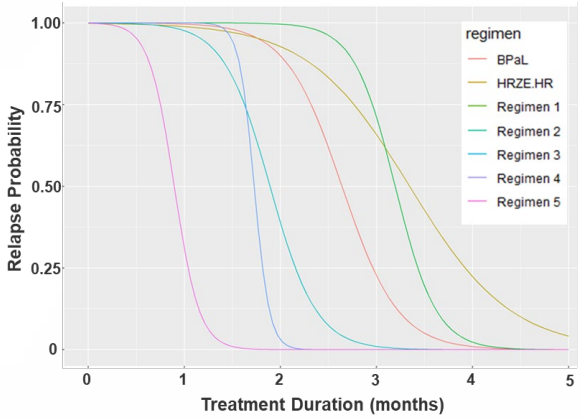
Simulation Process / Workflow

Determine study design:
- Regimens (real or hypothetical)
- Covariate values
- Sampling time points
- Number of animals

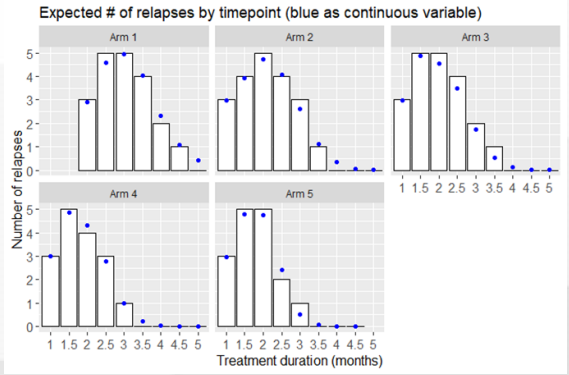


Slope,
Intercept

Randomly sample from model estimated distributions to assign mean parameter values for each virtual study (N = 1000)



Establish "true" relapse probability vs. time profiles



Simulate relapses in "virtual" mice

Statistical Model

A mixed effects logistic regression model was applied to account for inter-study variability, differential regimen response, and effects of study-level covariates

$$logit(p_{i,j,k}) = INT_{i,j} + SLP_{i,j} * (TIME_k - 2)$$

$$INT_{i,j} = INT_{TRT} + 1.386 \text{ x } (INOC_j - 3.29) + -1.505 \text{ x } I_{Sch7} + \sigma_{INT,j}$$

$$SLP_{i,j} = SLP_{TRT} + 0.497 * (BASE_j - 6.79) + \sigma_{SLP,j}$$

1000 studies were simulated by drawing etas from the variance-covariance matrix (σ_{INT}=1.209, σ_{SLP}=0.269, correlation of -0.72) , and assigning relapse by drawing from a random binomial distribution

Each simulated study was then pooled with historical data and the model re-estimated

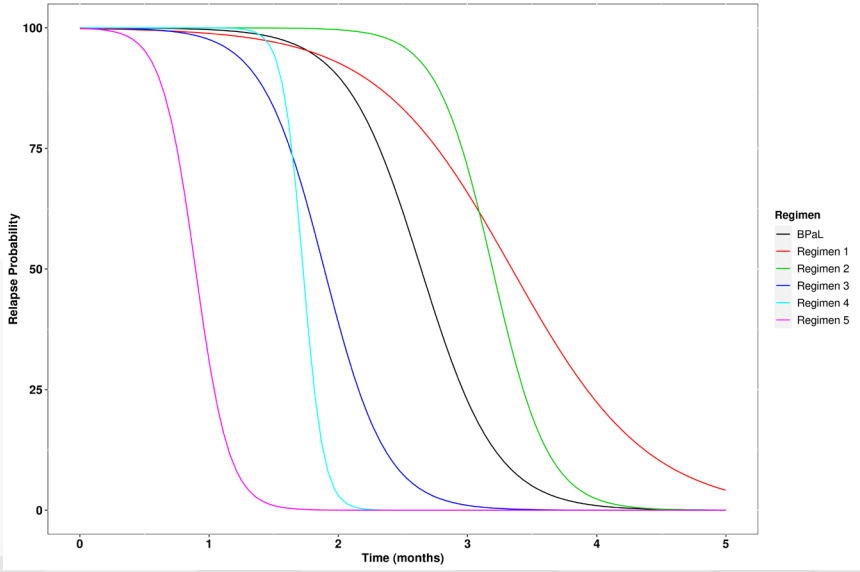
Note – Model is based on the Berg et al (2022) publication, with updates to include additional covariates

Simulation study – Regimen assumptions

Simulations were performed with the following inputs:

- Single BPaL control arm
- 7/7 dosing frequency

Regimen	Time in months to 5% Relapse (T5)	Time in months to 50% Relapse (T50)	Rank Order Based on T5 (1 = shortest duration)
BPaL	3.50	2.64	4
Regimen 1	4.90	3.34	6
Regimen 2	3.83	3.19	5
Regimen 3	2.60	1.89	3
Regimen 4	1.96	1.73	2
Regimen 5	1.28	0.89	1



Simulation Study - Design Attributes

The following major design attributes were evaluated:

- Number of time points
- Mice per time point
- Total mice

Covariates were fixed for inoculum (4.5 CFU/mL) and baseline bacterial burden (7.5 CFU/mL) according to real world targets for study

Simulations were done in 2 stages:

1. Preliminary simulations for general study design evaluation
2. Testing of additional designs to determine if mice could be better utilized and fit within logistical considerations

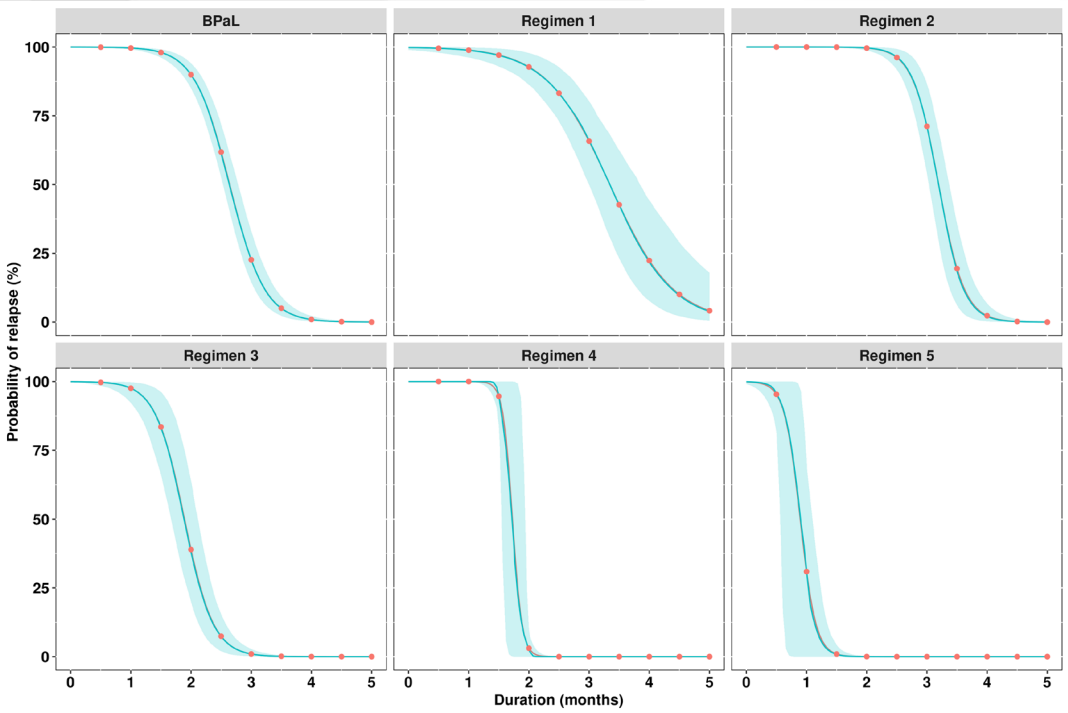
Simulation Study - Initial Study Designs

Design evaluation began with comparison of proposed study design relative to the "typical" design, a literature-based design* and some plausible modifications thereof, as well as a high-end benchmark

Design	Description	# Timepoints	Mice / timepoint	Relapse timepoints (Months of treatment)	Total mice / regimen
1	"Typical" RMM study design	4	15	1.0, 2.0, 3.0, 4.0	60
2	Erasmus design (based on Mourik et al.)	10	3	0.5, 1.0, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 4.5, 5.0	30
3	Weekly timepoints from 1 to 3 months plus biweekly to 4 months	12	3	0.5, 1.0, 1.25, 1.5, 1.75, 2.0, 2.25, 2.5, 2.75, 3.0, 3.5, 4.0	36
4	Weekly timepoints from 0.5 to 3 months plus biweekly to 5 months	15	3	0.5, 0.75, 1.0, 1.25, 1.5, 1.75, 2.0, 2.25, 2.5, 2.75, 3.0, 3.5, 4.0, 4.5, 5.0	45
5	Large sample size design with 5- 10x more mice (high-end benchmark)	10	30	0.5, 1.0, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 4.5, 5.0	300
6	**Original design to be evaluated**	**3**	**15**	**1, 1.5, 2**	**45**

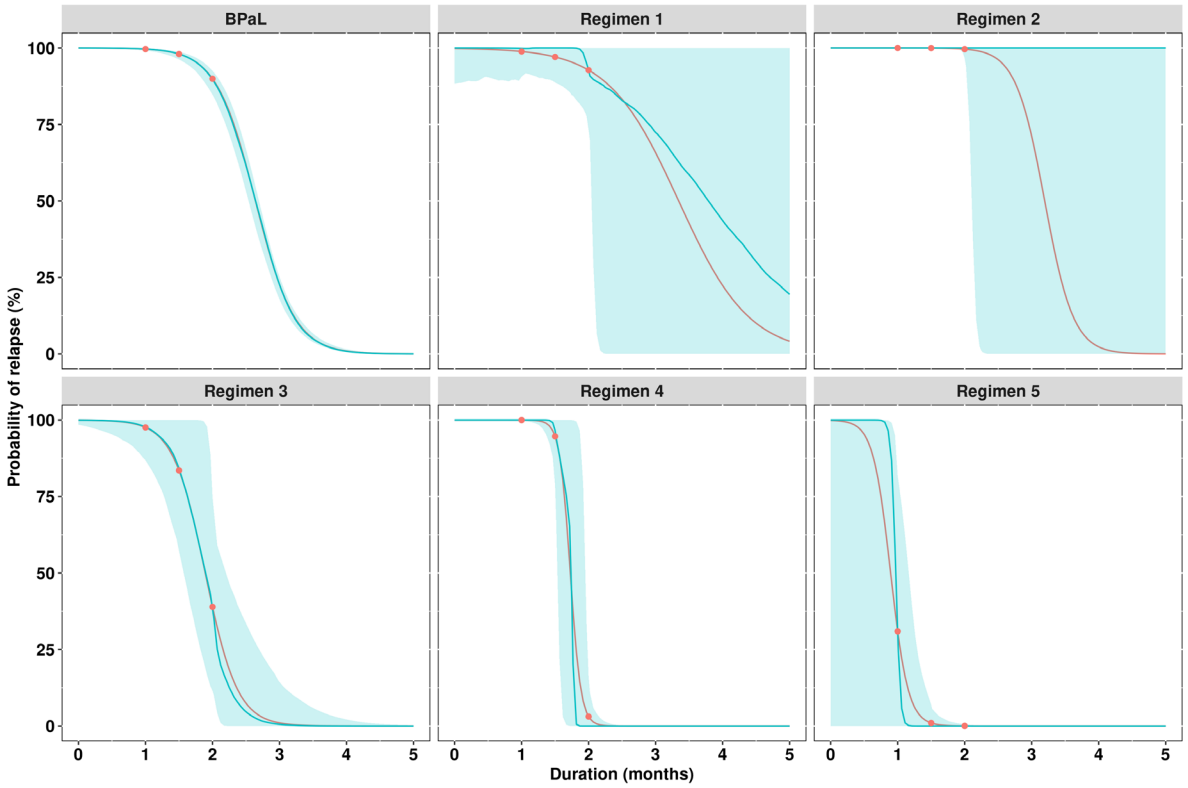
*Mourik et al. Sci Rep. 2018 Apr 9;8(1):5714. doi: 10.1038/s41598-018-24067-x.

Results for Selected Designs – High end benchmark



High end benchmark demonstrates the "best" that could be achieved with 5 – 10x more mice and increased timepoints

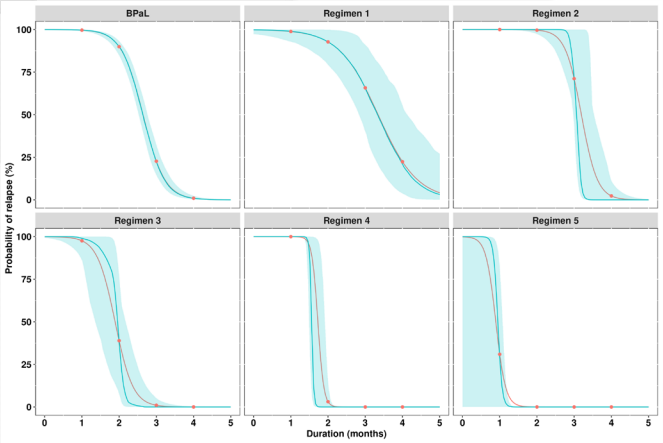
Results for Selected Designs – Original Design



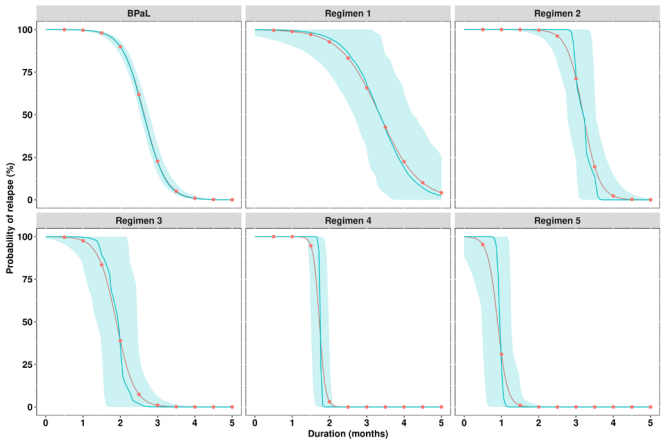
The partner's proposed design performed reasonably well for more efficacious regimens but not for regimens without a significant effect prior to 2 months due to there being minimal data to inform model estimates

Results for Selected Designs – Alternatives

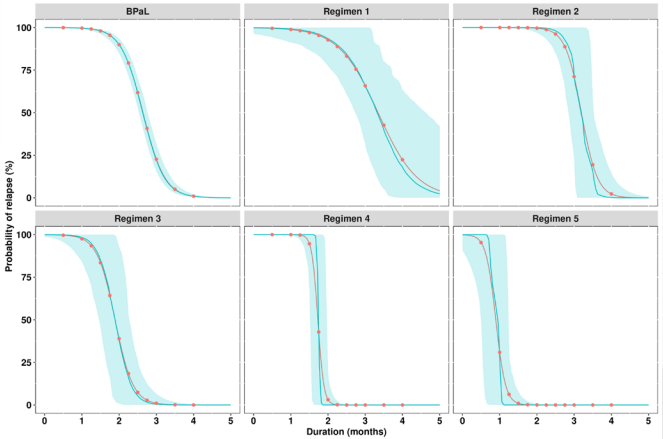
"Typical" RMM Design (4 timepoints, N=60/arm)



Erasmus Design (10 timepoints, N=30/arm)



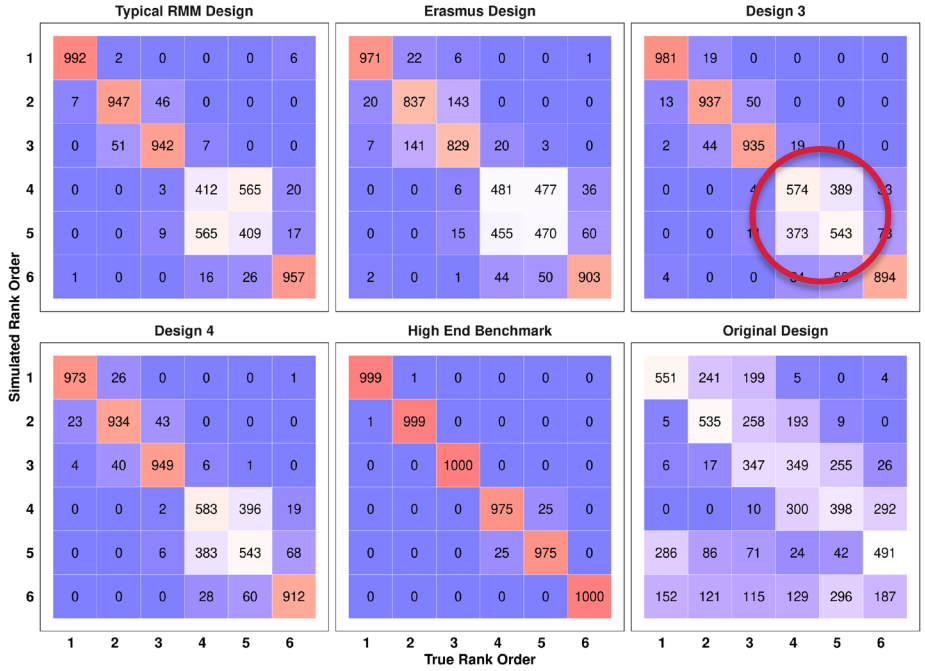
Design 3 (12 timepoints vs. Erasmus, N=36/arm)



More frequent sampling resulted in similar performance vs. the "Typical" design but with ~50% less mice. Slight improvements seen with modification of number / spacing of timepoints

Results – Rank order based on T_5

Assessment of regimen rank order based on T_5 indicates that Proposed design performs poorly compared to other designs, particularly as T_5 and T_{50} are greater than 1.5 months



Regimen	Rank Order Based on T5 (1 = shortest duration)	Time in months to 5% Relapse (T5)
Regimen 5	1	1.28
Regimen 4	2	1.96
Regimen 3	3	2.60
BPaL	4	3.50
Regimen 2	5	3.83
Regimen 1	6	4.90

Simulations suggest plausible alternatives have been identified, but they have difficulty differentiating rank order between regimens with T_5 within +/- 2 weeks of each other

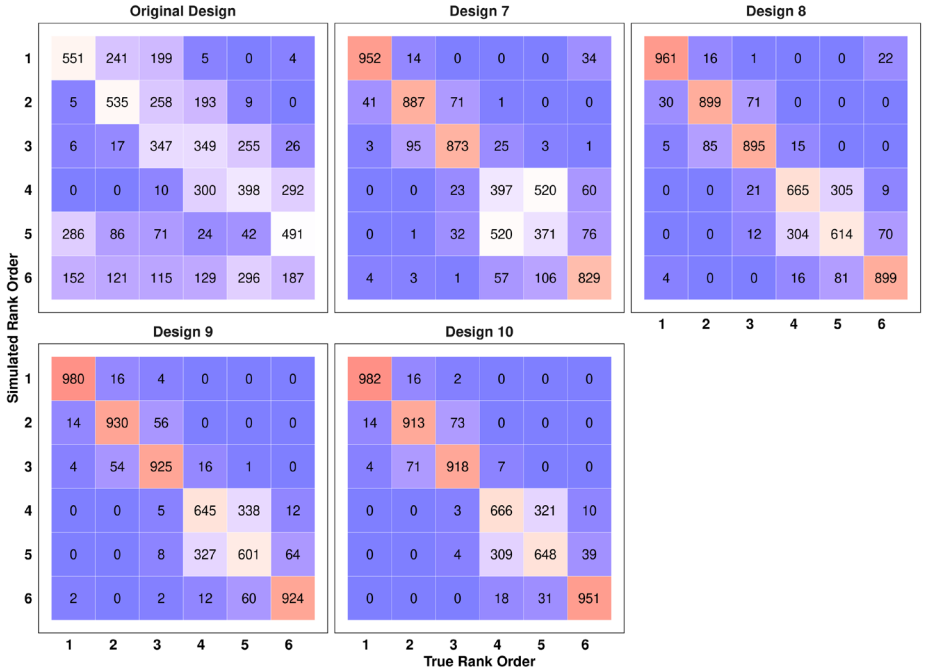
Additional Study Designs Simulated

Based on the initial simulation results, further alternative designs were explored to improve the discrimination between regimens, improve study feasibility, and better utilize available mice to maximize information

Design	Description	# Timepoints	Mice / timepoint	Relapse timepoints (Months of treatment)	Total mice / regimen
1	"Typical" RMM study design	4	15	1.0, 2.0, 3.0, 4.0	60
2	Erasmus Design (based on Mourik et al.)	10	3	0.5, 1.0, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 4.5, 5.0	30
3	Weekly timepoints from 1 to 3 months plus biweekly to 4 months	12	3	0.5, 1.0, 1.25, 1.5, 1.75, 2.0, 2.25, 2.5, 2.75, 3.0, 3.5, 4.0	36
4	Weekly timepoints from 0.5 to 3 months plus biweekly to 5 months	15	3	0.5, 0.75, 1.0, 1.25, 1.5, 1.75, 2.0, 2.25, 2.5, 2.75, 3.0, 3.5, 4.0, 4.5, 5.0	45
5	Large sample size design with 5- 10x more mice (high-end benchmark)	10	30	0.5, 1.0, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 4.5, 5.0	300
6	**Original design to be evaluated**	**3**	**15**	**1, 1.5, 2**	**45**
7	Biweekly timepoints from 1 to 3 months plus a 4 month timepoint	6	5	1, 1.5, 2, 2.5, 3, 4	30
8	Biweekly timepoints from 1 to 4 months	7	5	1, 1.5, 2, 2.5, 3, 3.5, 4	35
9	Biweekly timepoints from 0.5 to 4 months	8	5	0.5, 1, 1.5, 2, 2.5, 3, 3.5, 4	40
10	Biweekly timepoints from 0.5 to 4.5 months	9	5	0.5, 1, 1.5, 2, 2.5, 3, 3.5, 4, 4.5	45

Results – Rank order based on T_5 for proposed alternatives

Assessment of regimen rank order based on T_5 indicates that Designs 8 – 10 assigned the correct rank order for all regimens more than 60% of the time



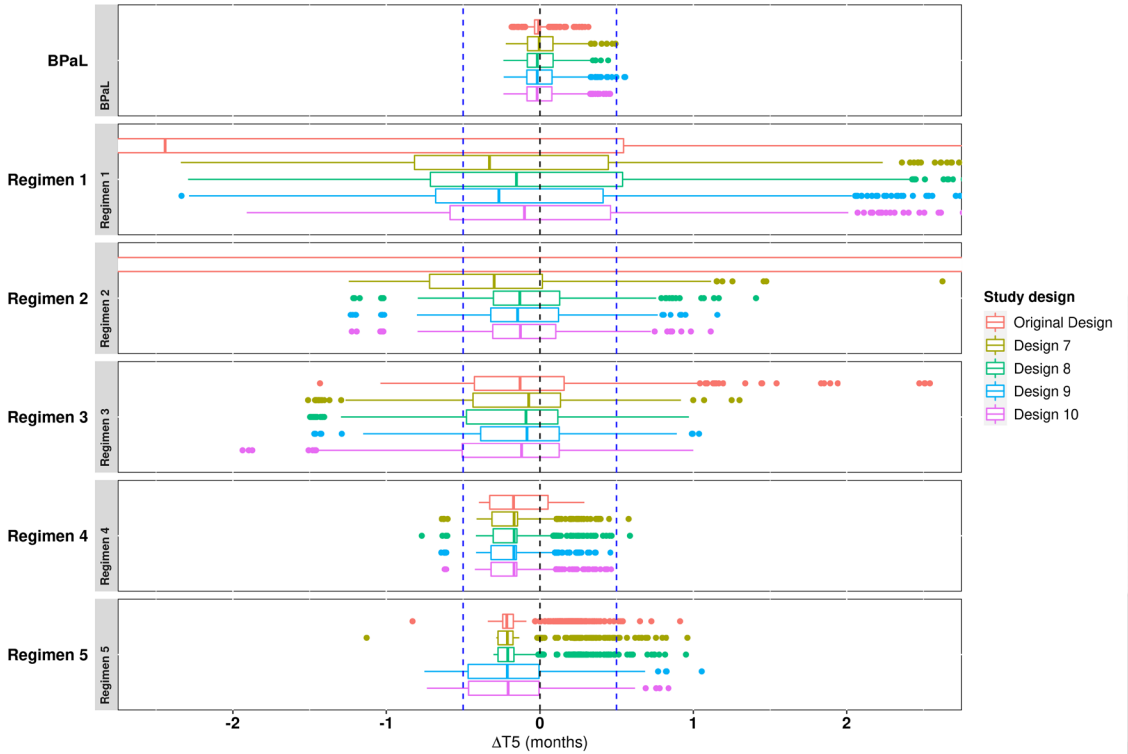
Regimen	Rank Order Based on T5 (1 = shortest duration)	Time in months to 5% Relapse (T5)
Regimen 5	1	1.28
Regimen 4	2	1.96
Regimen 3	3	2.60
BPaL	**4**	**3.50**
Regimen 2	**5**	**3.83**
Regimen 1	6	4.90

Results – T_5 Bias plot for proposed alternatives

The alternate designs (Designs 7 – 10) showed improvements in T_5 bias / precision relative to the proposed design

- Consistent negative median bias across all hypothetical regimens across all designs, suggesting tendency towards underestimation of T_5 by ~1 week
  - Bias was only resolved in the high-end benchmark study design, suggesting that a significantly increased sample size would be required to mitigate this minor bias

Bias plot for T_5 by regimen / study design



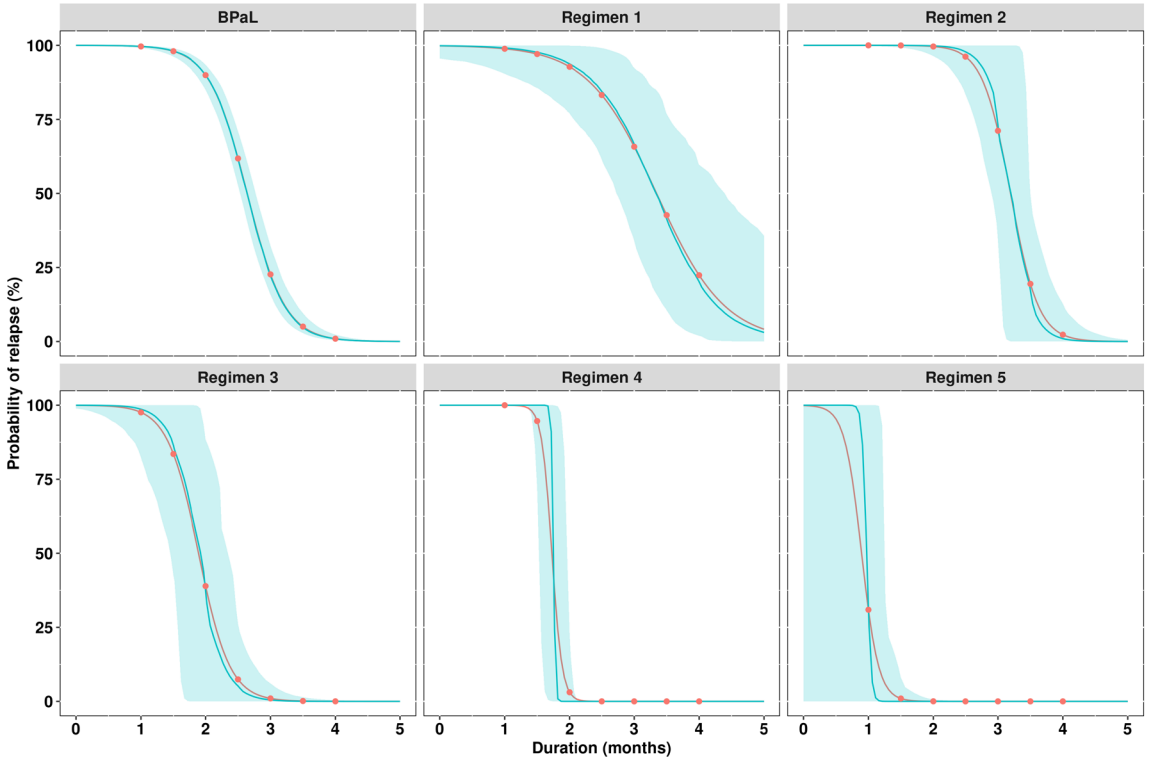
Outcome – Team Recommendation

A proposed design (Design 8) was recommended to the partner as it exhibited the best tradeoff between number of mice and the number of timepoints

- The proposed design performed much better than the original design and shows some improvements over the “typical” design which traded off a much larger sample size for fewer timepoints
- The proposed design also exhibited better performance over Design 2 (Erasmus) and Design 7, suggesting a benefit in increasing the sample size from 30 to 35 mice
  - Gains appeared to diminish with increasing sample size as it exhibited similar performance vs. Designs 9 and 10

Design	Description	# Timepoints	Mice / timepoint	Relapse timepoints (Months of treatment)	Total mice / regimen
1	“Typical” RMM study design	4	15	1.0, 2.0, 3.0, 4.0	60
2	Erasmus Design (based on Mourik et al.)	10	3	0.5, 1.0, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 4.5, 5.0	30
6	Original design to be evaluated	3	15	1, 1.5, 2	45
7	Biweekly timepoints from 1 to 3 months plus a 4 month timepoint	6	5	1, 1.5, 2, 2.5, 3, 4	30
8	**Biweekly timepoints from 1 to 4 months**	**7**	**5**	**1, 1.5, 2, 2.5, 3, 3.5, 4**	**35**
9	Biweekly timepoints from 0.5 to 4 months	8	5	0.5, 1, 1.5, 2, 2.5, 3, 3.5, 4	40
10	Biweekly timepoints from 0.5 to 4.5 months	9	5	0.5, 1, 1.5, 2, 2.5, 3, 3.5, 4, 4.5	45

Outcome – Recommended Design (Design 8)



The recommended design was able to better capture most regimens including those without significant effect prior to 2 months

Summary / Conclusions

Through application model-informed drug development tools and techniques, we were able to critically assess a set of key non-clinical studies and recommend meaningful improvements to the study design

- A series of RMM studies were proposed by a partner organization
- The performance of a proposed design compared to alternative designs was explored through a comprehensive simulation / re-estimation exercise
- Additional refinements were evaluated, resulting in an improved design recommendation which was taken forward by the partner organization

By switching to the recommended study design, the partner will obtain significantly more informative data from the experiment to inform regimen development decisions. Moreover, the recommended design decreases the number of animals utilized in a terminal study by over 20% versus the originally proposed design, promoting better animal stewardship in R&D efforts

Acknowledgements

- Alexander Berg and Jessica Roberts (Simulations Plus)
- David Hermann and Debra Hanna (Gates Foundation)
- Lab of Eric Nuermberger (Johns Hopkins)
- Lab of Anne Lenaerts (Colorado State)
- Critical Path Institute
- CPTR Initiative member organizations

SimulationsPlus

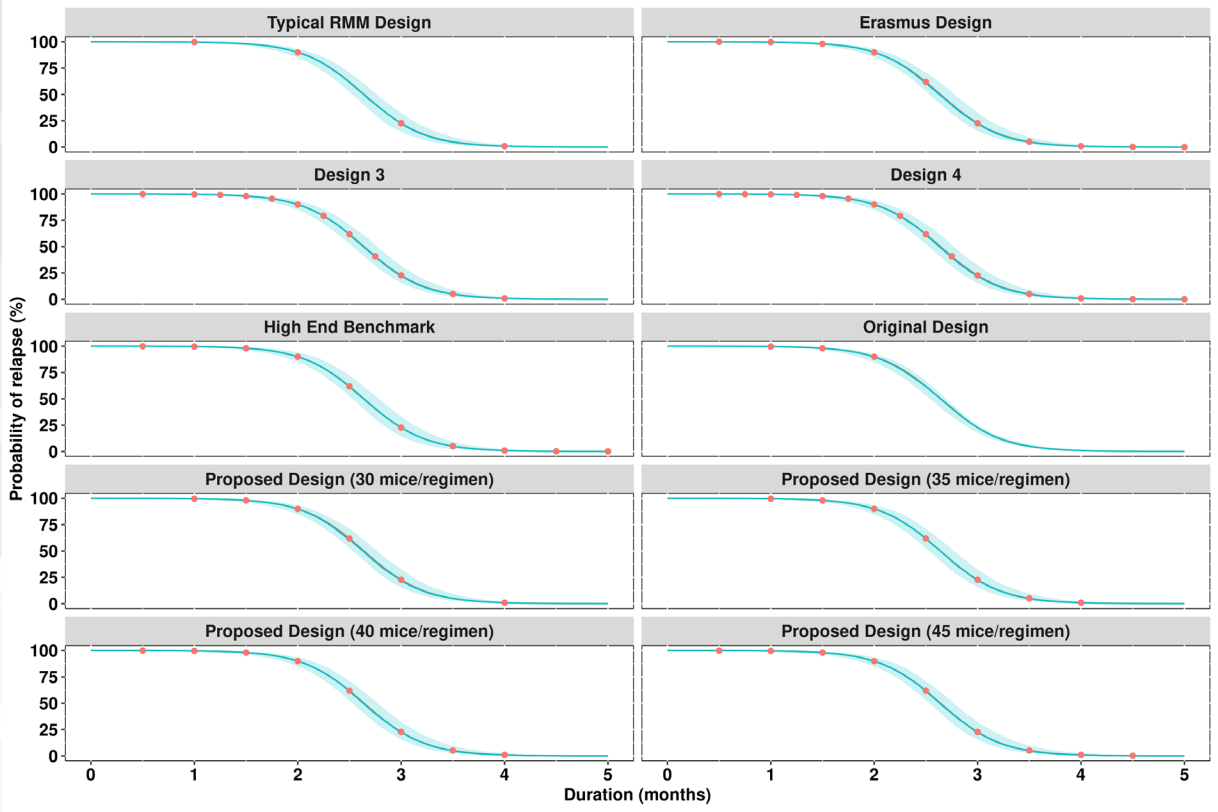
MIDD+22

Model Informed Drug Development

Q&A

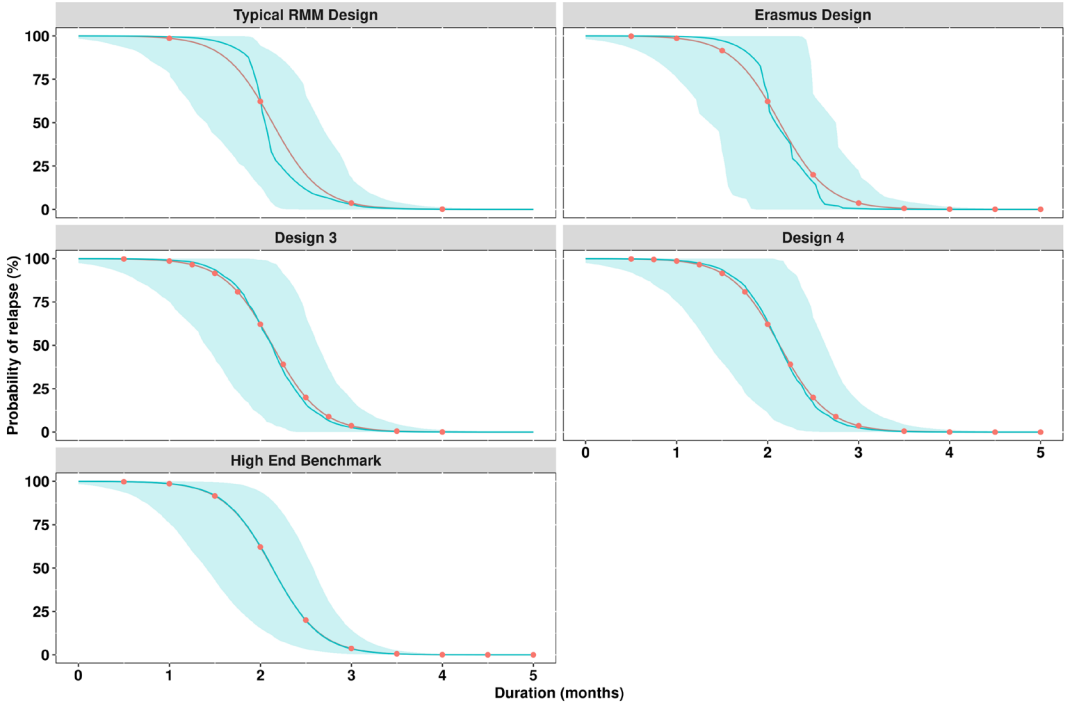
Questions & Answers

Simulation Results – Control Arms with Historical Data



Pooling simulated study data with historic data allows for estimation of control in proposed design and gives an anchor to aid in estimation of study level random effects

Simulation Results – Control Arms without Historical Data



Initial testing of simulation-estimation with out historical data showed wide confidence intervals for controls

Simulation Results – Performance for Time to 5% Relapse

Negative bias of ~ 1 week is consistent across designs (except high end benchmark)

- Extreme bias and wide confidence intervals seen for original design for regimens 1 and 2 where T_5 is > 3 months
- Increased bias in regimen 5 for original design

Original designs performs the worst for the given regimens compared to the other designs

Bias plot for T_5 by regimen / study design