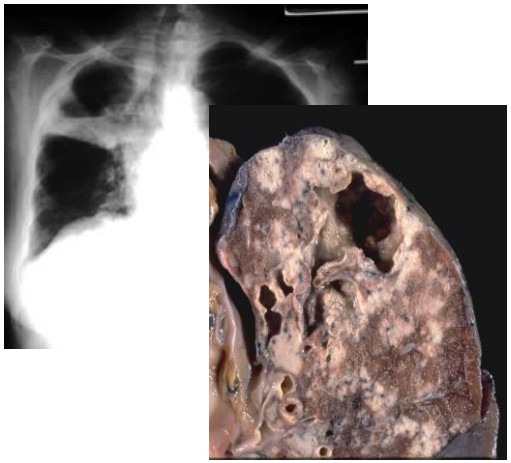
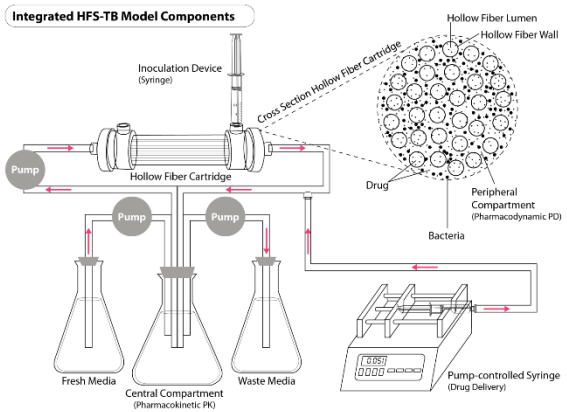
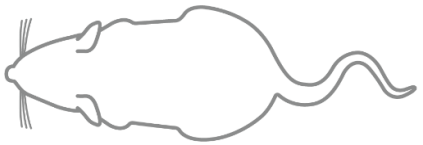
Examining the Predictive Accuracy of Sterilizing Mouse Efficacy Models

Eric Nuermberger, MD
Center for TB Research, Johns Hopkins University

March 20, 2107

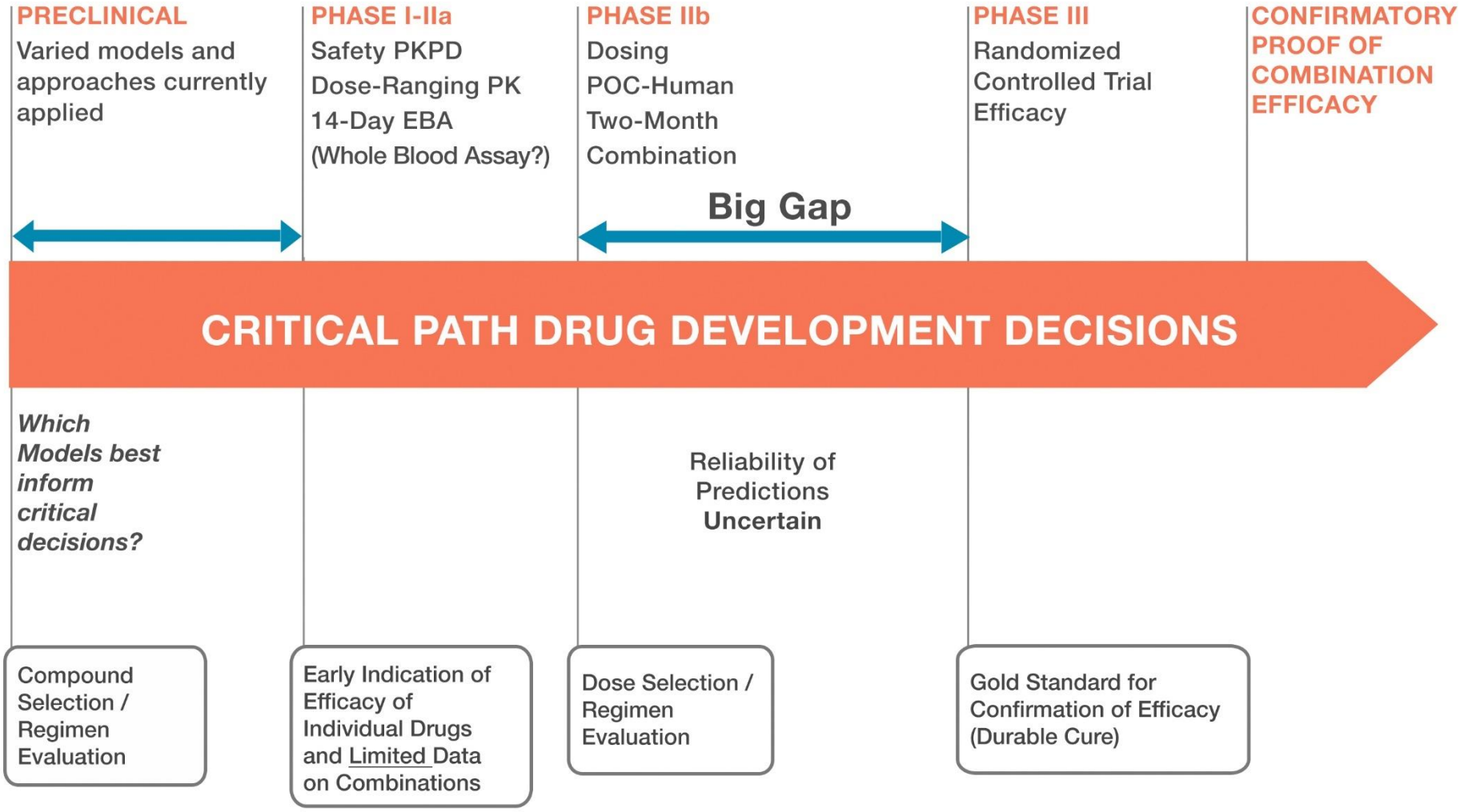






Current TB regimen development

Risk of late-stage attrition



Pre-Clinical and Clinical Sciences Workgroup (PCS-WG) Mission & Goals

Mission

Develop and/or validate tools and innovative approaches to address pre-clinical issues including *in vitro* and *in vivo* efficacy, PKPD analyses using appropriate biomarkers, drug safety, metabolism, DDI, etc. These tools may be submitted to regulatory authorities for regulatory review and/or qualification as appropriate.

Early goal related to pre-clinical *in vitro* and *in vivo* models

Evaluate the evidence base and develop criteria for evaluating the utility of various preclinical models to inform and test new drug regimens.

Early Evidence

Landscape analysis* identified HFS-TB as having an appropriate data inventory to assess predictive accuracy of a preclinical model for clinical outcomes.

*Gumbo et al, JID 2015; 211(S3):S83

Early success

EMA qualification opinion on the HFS-TB June 26, 2014

- HFS-TB qualified for use in drug development programs ***as additional and complementary tool***
- HFS-TB can be used in regulatory submissions, esp. for informed design and interpretation of clinical studies
- HFS-TB is recommended to be useful as follows:
  - To provide preliminary proof of concept for developing a specific drug or combination to treat tuberculosis
  - To select the pharmacodynamic target (e.g. $T_{>MIC}$, AUC/MIC)
  - To provide data to support PK/PD analyses leading to initial dose selection for non-clinical and clinical studies
  - To assist in confirming dose regimens for later clinical trials taking into account human PK data and exposure-response relationships

Evaluation of *in vivo* models

"Correlations between drug concentration and pathogen survival that are based on in vitro models cannot be expected to reiterate all aspects of in vivo antimycobacterial treatment."

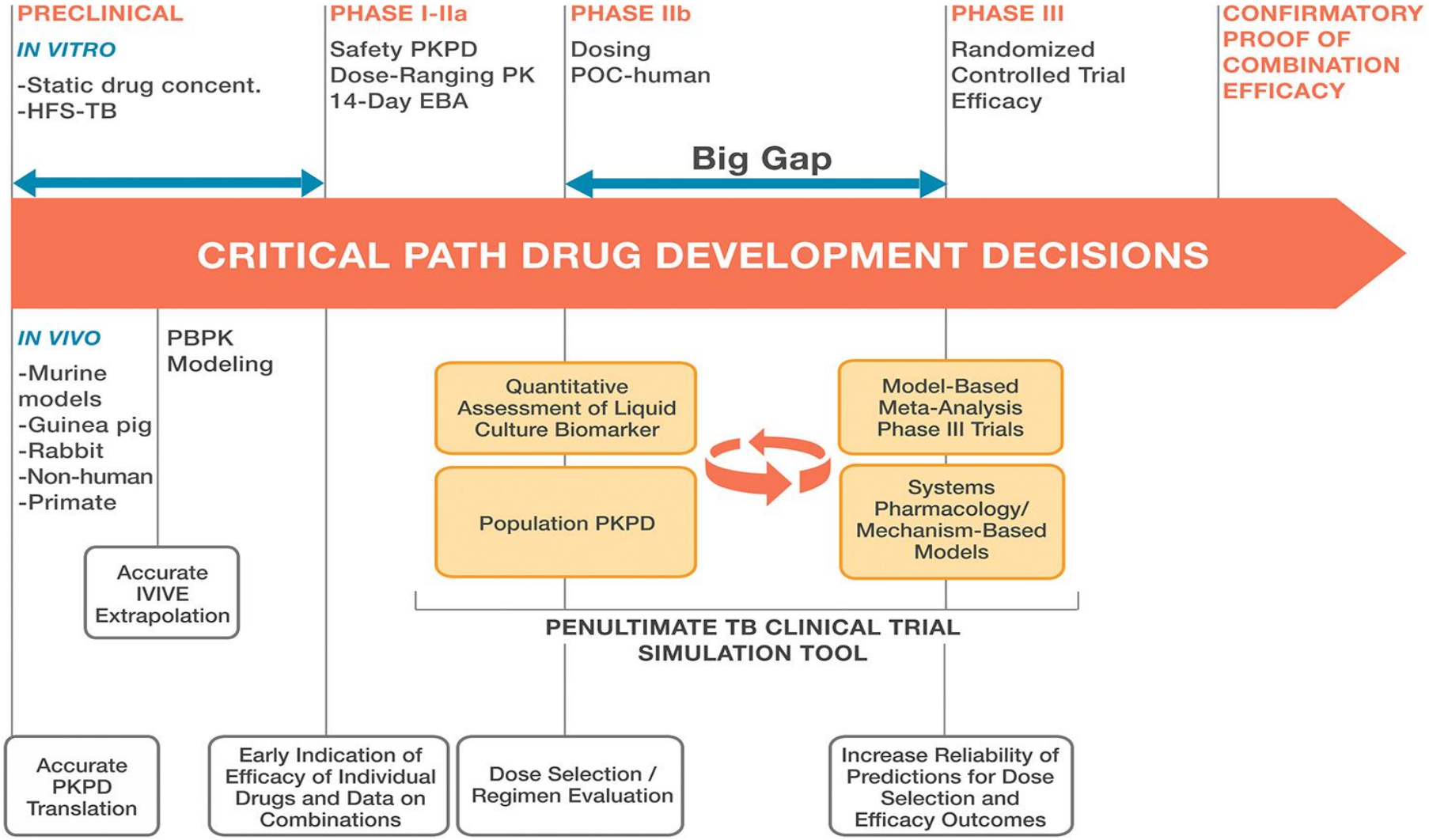
Chilukuri et al, CID 2015; 61(S1):S32

Advantages of *in vivo* models

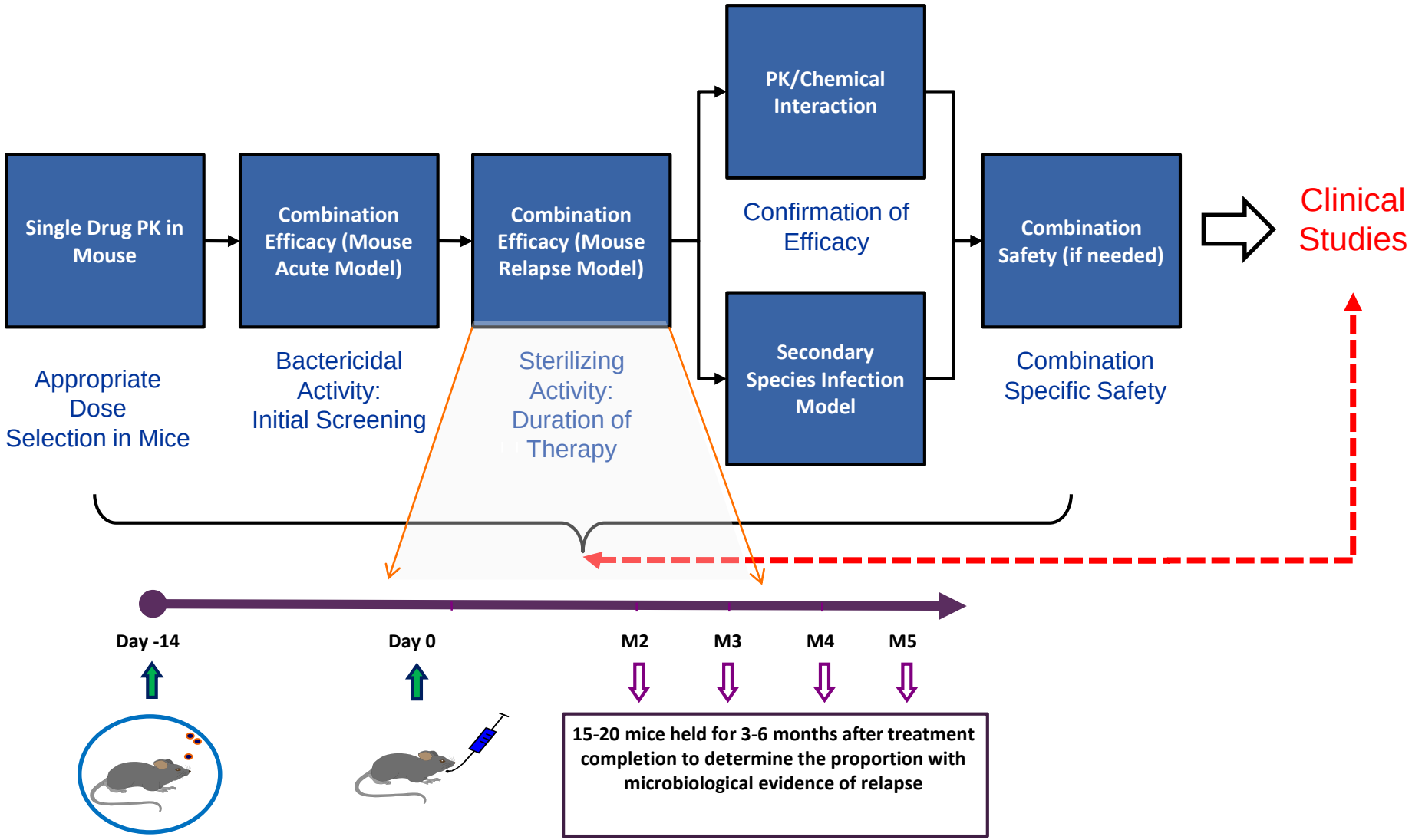
- Better reflect the phenotypic heterogeneity in bacterial populations as determined by host-pathogen interactions, including development of tissue pathology
- Present complexities of drug distribution to, and action at, various sites of infection

Improving TB regimen development

Decreasing risk of attrition



Mouse model of sterilizing activity



Evaluating the sterilizing mouse model

Rationale

- Past and present role in TB regimen development
  - track record in forecasting treatment-shortening potential of RIF, PZA
  - relapse endpoint considered closest correlate of current phase 3 endpoint
- Amount of available data on regimens evaluated in clinical trials
- Does not preclude evaluation of other models

Evaluating the sterilizing mouse model

Rationale

- Past and present role in TB regimen development
  - track record in forecasting treatment-shortening potential of RIF, PZA
  - relapse endpoint considered closest correlate of current phase 3 endpoint
- Amount of available data on regimens evaluated in clinical trials
- Does not preclude evaluation of other models

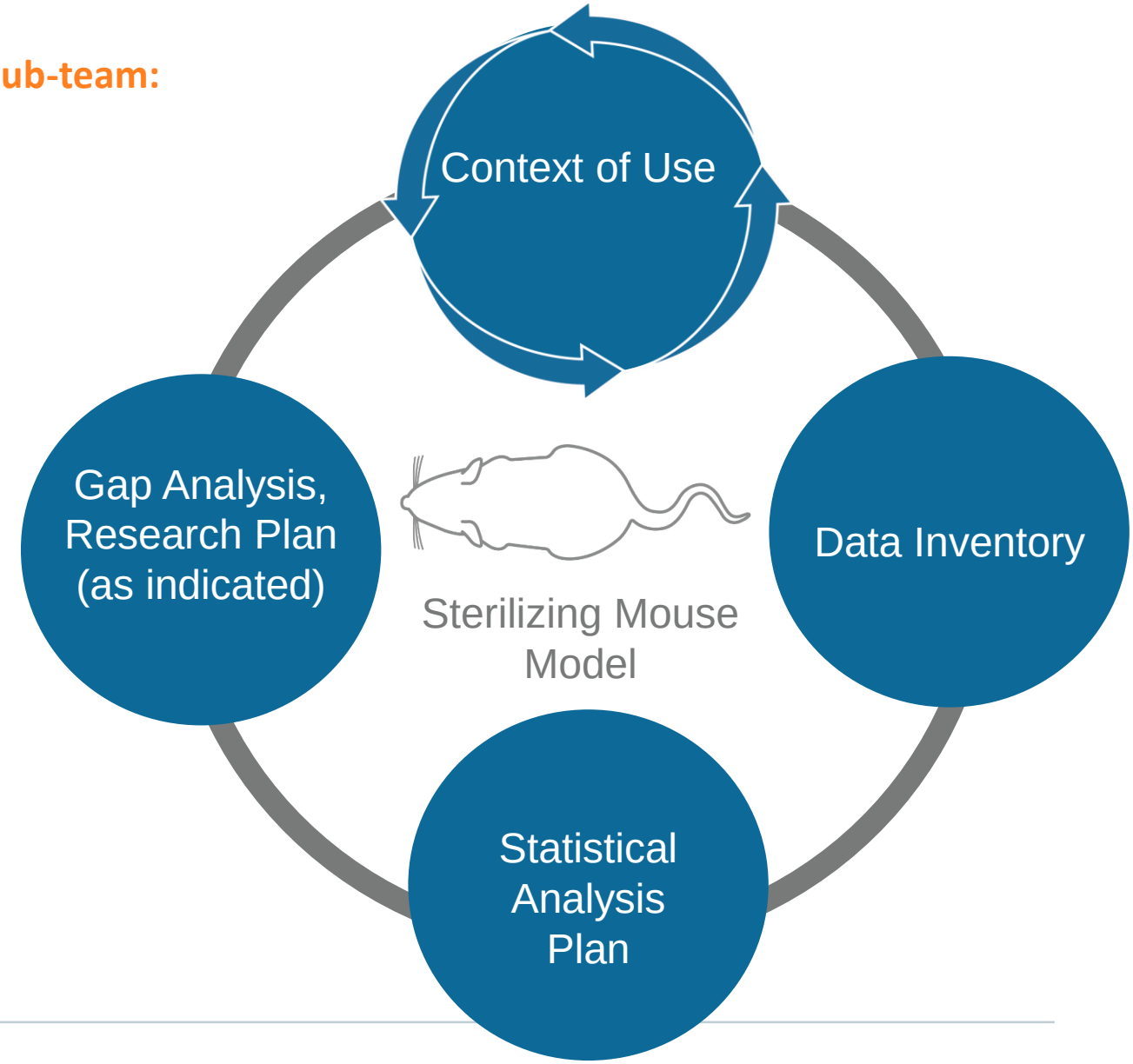
General Aim

- **Quantify the predictive accuracy** of mouse TB efficacy models **to rank order regimens** and **estimate the effective treatment duration**, by evaluating the proportional and absolute differences in the treatment durations of test and control regimens required to produce the same relapse outcome

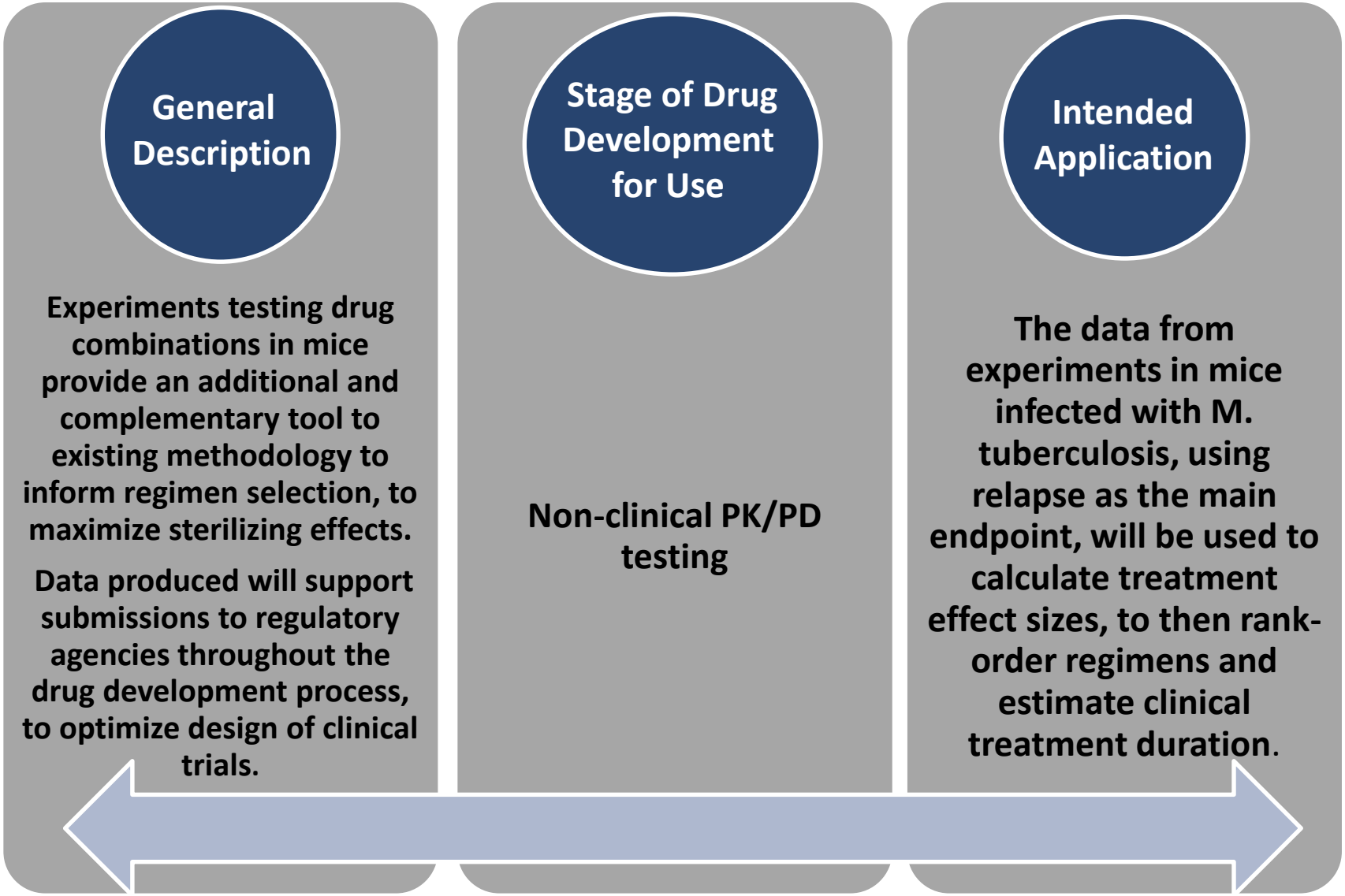
Workplan for evidence-based evaluation of sterilizing mouse model

CPTR PCS-WG Mouse Model Sub-team:

Dr. Dakshina Chilukuri
Dr. Geraint Davies
Dr. Geo Derimanov
Dr. Nader Fotouhi
Dr. Tawanda Gumbo
Dr. Debra Hanna
Dr. Karen Lacourciere
Dr. Barbara Laughon
Dr. Anne Lenaerts
Dr. Owen McMaster
Dr. Nader Fotouhi
Dr. Eric Nuermberger
Dr. Klaus Romero
Dr. Rada Savic
Dr. Christine Sizemore
Dr. Peter Warner
Lindsay Lehmann



COU Scenario 1: Treatment Duration and Rank-Ordering of Regimens



Data inventory

- Focus first on mouse strains other than C3HeB/FeJ ("Kramnik")
- Inventory identified a variety of relapse-based pre-clinical studies with corresponding clinical trial outcomes data

Test regimen intervention	Regimen comparison	# of expts
Combining INH+STR	HS vs. H or S monotherapy	1
Shortening duration of INH+STR	6HS vs. 18HS	1
Adding RIF to INH+STR or INH+EMB+PZA	HR (or HRS or HREZ) vs. HS (or HEZ)	4
Adding STR to INH+RIF	HRS vs. HR	1
Adding PZA to INH+RIF (±STR/EMB)	HRZ (or HRSZ or HREZ) vs. HR (or HRS or HRE)	4
Shortening duration of PZA	2HREZ/4RH vs. 6HREZ	1
Increasing dose of RIF	High-dose R plus HEZ vs. HREZ	2
Extending dosing interval of 1st-line Rx	HREZ (2/7) vs. HREZ (daily)	1
Replacing EMB with MXF	HRMZ vs. HRZ(E)	3
Replacing INH with MXF	MRZ(E) vs. HRZ(E)	10
Replacing RIF with RPT	HPZ(E) vs. HRZ(E)	7
Replacing RIF+EMB with RPT+MXF	HPMZ vs. HRZ	3
Replacing RIF with RPT and extending dosing interval (in continuation phase)	HP(1/7) cont phase vs. HR(2/7)	2
Comparing INH+RIF+PZA+EMB with PMD+MXF+PZA	PaMZ vs. HRZ(E)	8

Proposed statistical analysis plan

The following analysis approaches are proposed, with the objective of rank ordering specific regimens and estimating their clinical treatment duration:

A. Logistic regression to determine predictors of the proportional and absolute change in the treatment duration required to achieve the same probability of relapse, compared between control and test regimens.

B. Parametric time-to-event analysis to determine predictors of the time-varying probability of achieving the same proportion of relapses over time, evaluating control and test regimens.

C. Non-linear meta-regression analysis to determine specific interpretable parameters for specific proportional and absolute changes in treatment duration necessary to prevent relapses over time, evaluating control and test regimens.

Summary points

- An initial step to address the “translational gap” is to learn what data from what models analyzed in what way best inform key trial design decisions.
- Evidence-based validation of pre-clinical models is important:
  - to confidently place preclinical models on the critical development path,
  - to increase the efficiency of regulatory interactions,
  - to set a precedent for objective, data-driven processes to apply to other models (e.g., C3HeB/FeJ mouse, marmoset), and
  - to identify gaps in knowledge & in existing tools to drive future research.
- Evaluation of sterilizing mouse models is the appropriate first step for *in vivo* models, with other models to follow

Acknowledgements

CPTR PCS-WG Mouse Model Sub-team:

Dr. Nicole Ammerman (Johns Hopkins University)

Dr. Dakshina Chilukuri (US Food & Drug Administration)

Dr. Geraint Davies (University of Liverpool)

Dr. Geo Derimanov (Glaxo Smith Kline)

Dr. Nader Fotouhi (Global Alliance for TB Drug Development)

Dr. Tawanda Gumbo (Baylor University)

Dr. Debra Hanna (Critical Path Institute)

Dr. Karen Lacourciere

Dr. Barbara Laughon (National Institutes of Health)

Dr. Anne Lenaerts (Colorado St. University)

Dr. Owen McMaster (US Food & Drug Administration)

Dr. Khis Mdluli (Global Alliance for TB Drug Development)

Dr. Eric Nuermberger (Johns Hopkins University)

Dr. Klaus Romero (Critical Path Institute)

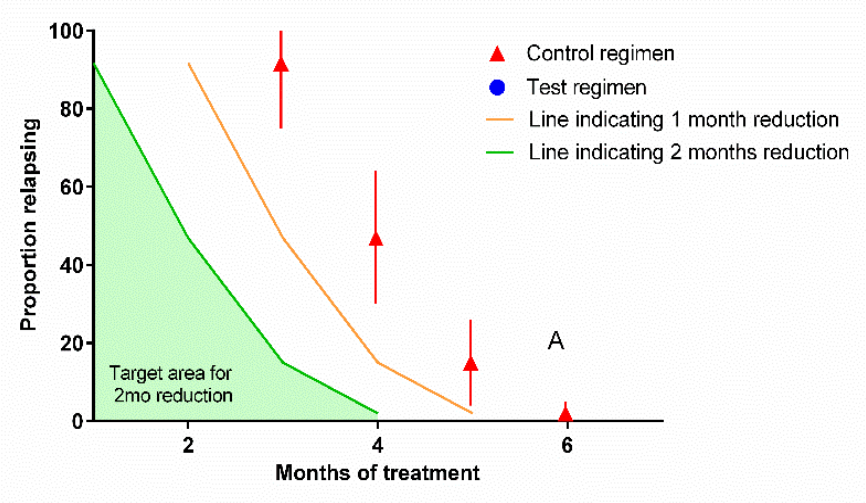
Dr. Rada Savic (University of California-San Francisco)

Dr. Christine Sizemore (National Institutes of Health)

Dr. Peter Warner (Bill & Melinda Gates Foundation)

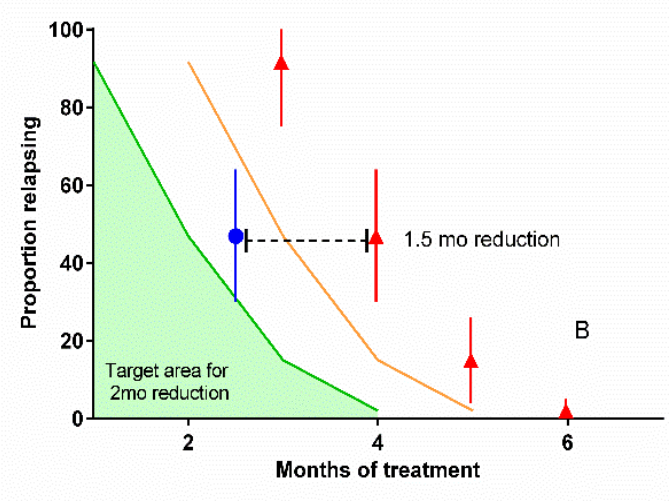
Lindsay Lehmann (Critical Path Institute)

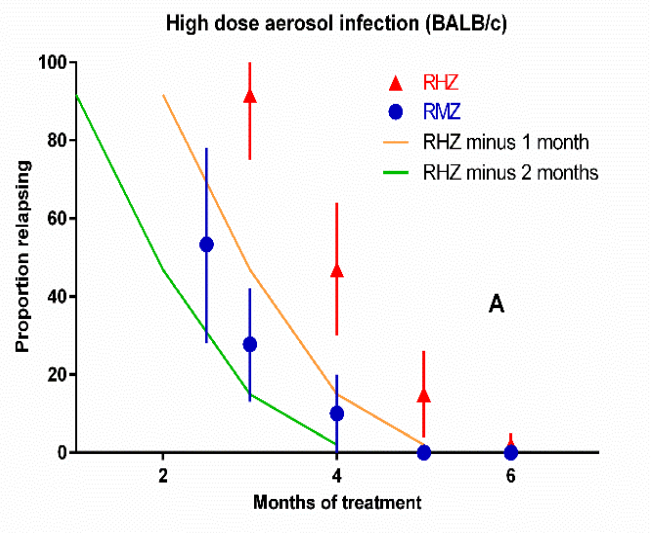
Estimating treatment duration



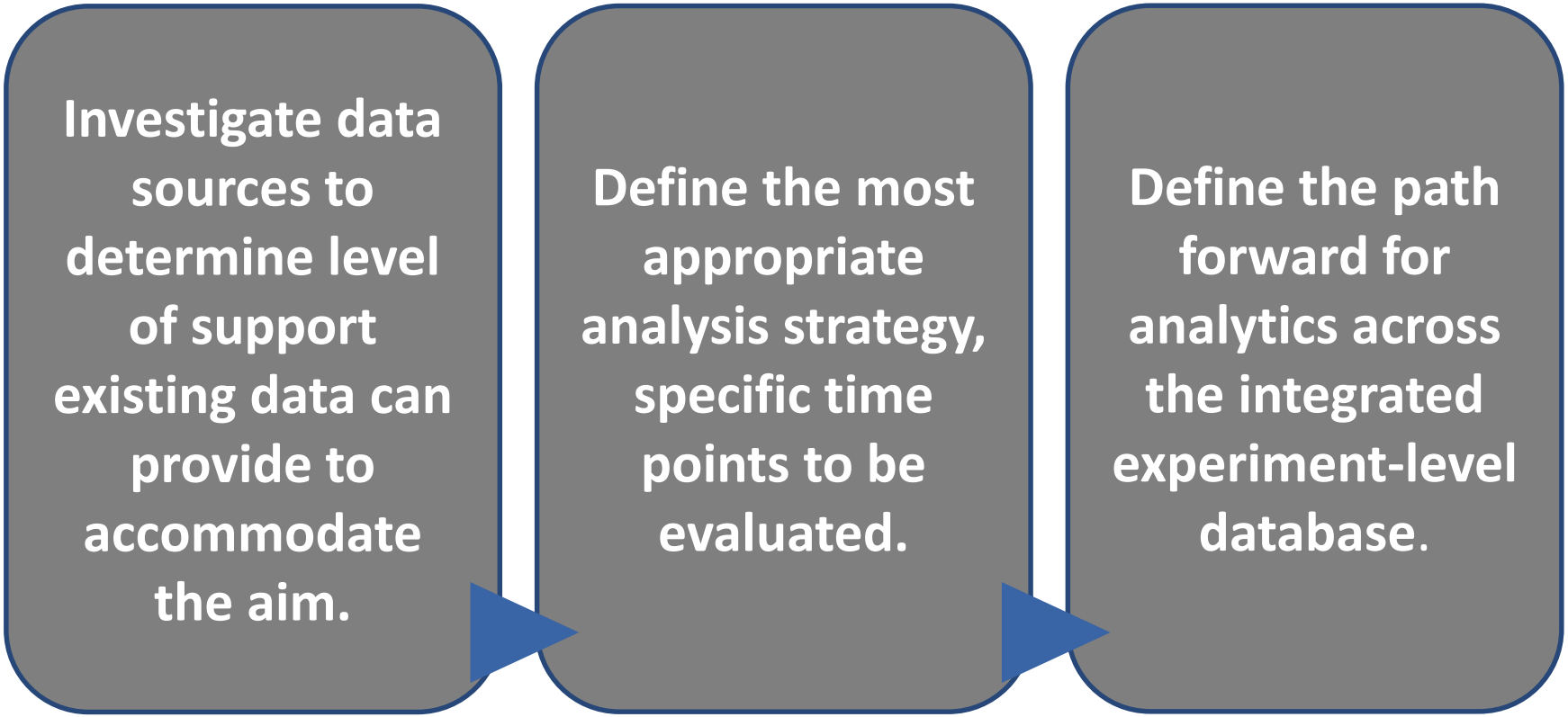
Effect size may be measured as difference in time to event (e.g., 50% of mice cured).

Both absolute and proportional effect sizes will be considered.

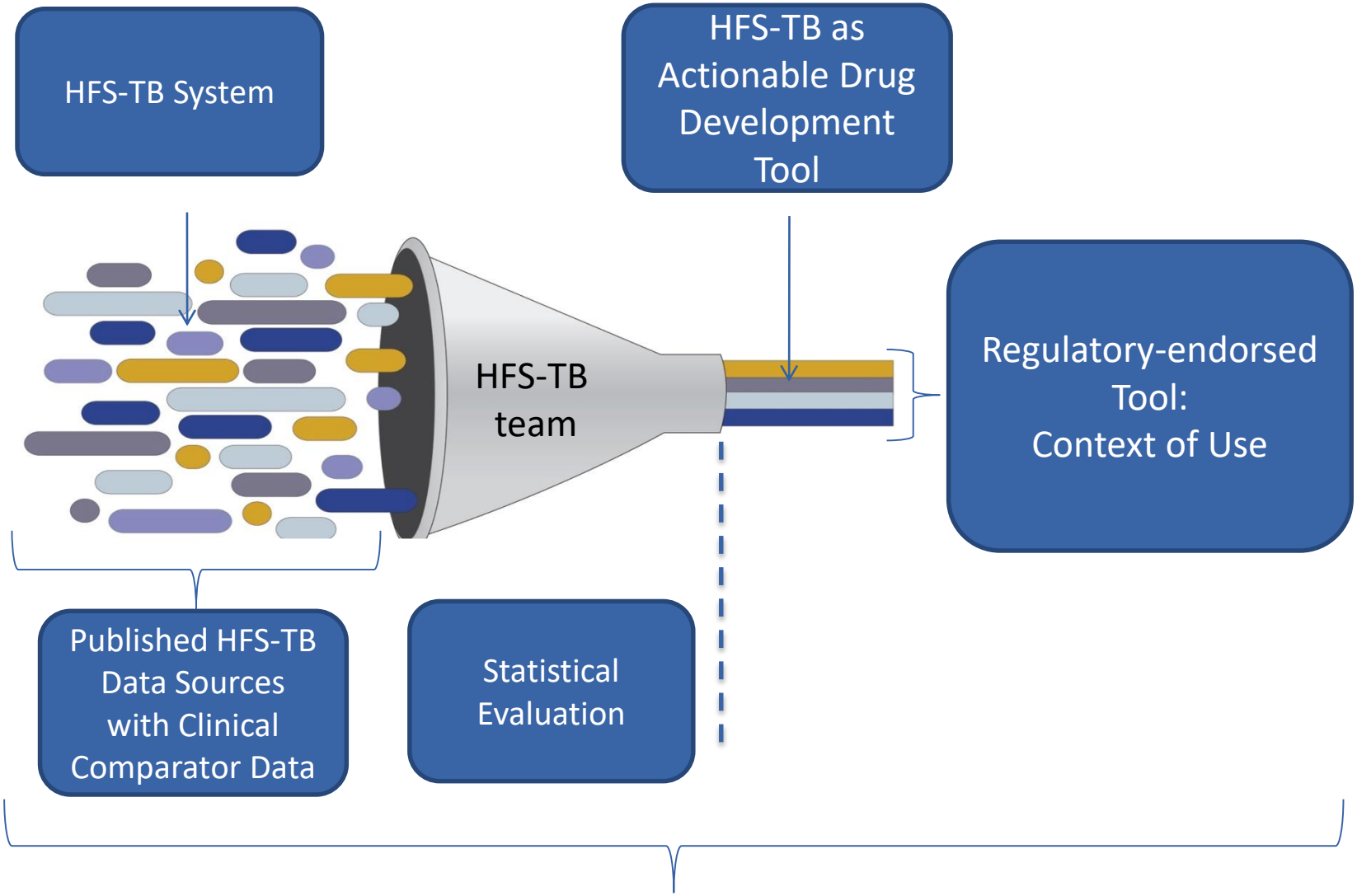




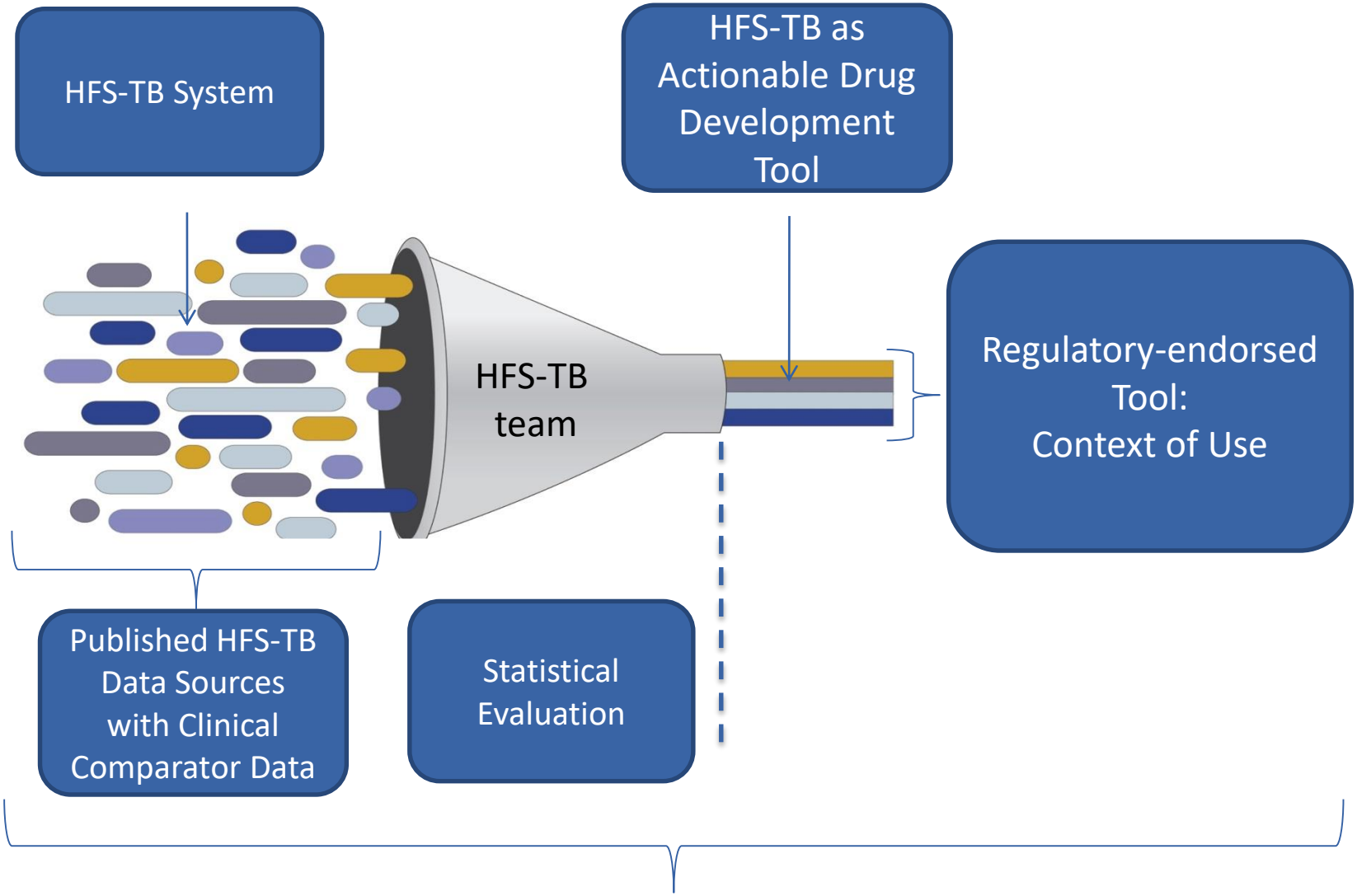
Development of statistical analysis plan



Pre-clinical DDT: CPTR HFS-TB Example



Pre-clinical DDT: CPTR HFS-TB Example



Context of use

- **Scenario 1:** Rank-ordering of regimens and estimation of treatment duration
- General description: The data from experiments *testing drug combinations* in mice infected with *M. tuberculosis* provide an *additional and complementary tool* to existing methodology *to inform regimen selection*, to maximize sterilizing effects. These data will support submissions to regulatory agencies throughout the drug development process for an anti-TB regimen, to optimize design of clinical studies.
- Stage of Drug Development for Use: Non-clinical PKPD testing.
- Intended Application: The data from experiments using mice infected with *M. tuberculosis*, using relapse as the main endpoint, will be used to calculate treatment effect magnitudes, to then rank-order regimens and predict clinical treatment duration.

Data Analysis Methods

Analysis 1: Descriptive Correlations

Analysis 2: Predictive Accuracy or Forecasting

- 2a: Correct ranking of PK/PD indices relevant to dose scheduling
- 2b: Accuracy in generating or refuting hypotheses with relevance to therapeutic strategies
- **2c: Quantitative accuracy in forecasting PK/PD indices relevant to dose scheduling, dose selection, and breakpoints**
  - ➢ *Weighted by clinical study quality score and number of patients in study*

Studies Identified by Searches

Literature Search A: 26 HFS-TB studies (12 combination studies, 10 monotherapy, 4 Monte Carlo simulations)

Literature Search B: 17 TB clinical studies, published prior to HFS-TB studies; quality of evidence score of 1 in 15/17

Literature Search C: 20 TB clinical studies, published at least six months after HFS-TB studies; quality of evidence of 1 or 2 in 11/20

- ***Weighting reflected clinical study quality score***

Predictive Accuracy Approach

- Error (E) was defined as the observed results in a clinical study at time T, minus the predicted value P:

$$E = T - P$$

- For a number of trials or experiments *i* of up to *n*, this takes the form of the mean absolute percentage error (MAPE), which is given by:

$$MAPE = \frac{1}{n} * \left[\sum_{i=1}^{n} \left| \frac{T_i - P_i}{T_i} \right| * 100 \right]$$

- Accuracy (A) was defined as:

$$A = 100\% - MAPE$$

- Bias (B) was defined as:

$$B = \sum_{i=1}^{n} (T_i - P_i)/n$$

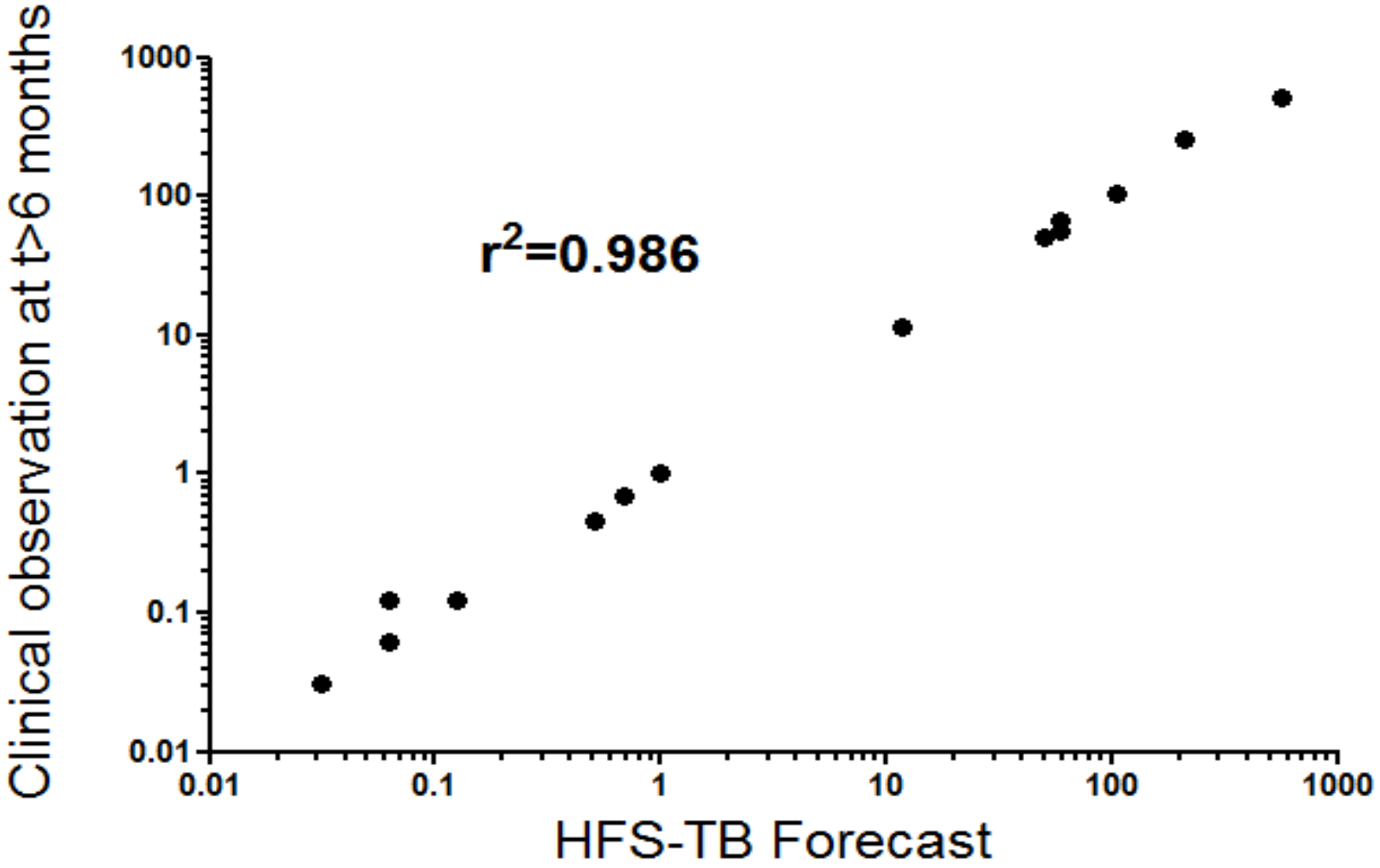
Table 4. Accuracy of the Hollow Fiber System Model of Tuberculosis in Predicting Quantitative Therapeutic Indices in Patients

Drug	Parameter	HFS-TB	Clinical Observation	No. of Patients	Weighting %	Weighted Accuracy	Weighted Bias (%)
Pyrazinamide [13, 40]	Optimal AUC/MIC	209	258	142	13.4	10.9	2.6
Isoniazid [12, 40]	Optimal AUC/MIC	567	520	142	13.4	12.2	–1.2
Ethambutol [21, 41]	Optimal peak/MIC	0.51	0.46	59	11.2	9.9	–1.2
Moxifloxacin [5, 42]	AUC/MIC at dose of 400 mg/d	59	66	9	0.9	0.8	0.1
Moxifloxacin [5, 45, 46]	AUC/MIC at dose of 400 mg/d	59	56	9	0.4	0.4	–0.0
Moxifloxacin [5, 47]	Optimal AUC/MIC	106	106	61	11.5	11.5	0
Pyrazinamide [13, 41]	Optimal AUC/MIC	11.7	11.3	59	11.2	10.8	–0.4
Rifampin [18, 43]	Breakpoint MIC, mg/L	0.0625	0.125	36	3.4	1.7	1.7
Rifampin [18, 48]	Breakpoint MIC, mg/L	0.0625	0.0625	52	2.2	2.5	0
Isoniazid [18, 43]	Lower resistance breakpoint MIC, mg/L	0.0312	0.0312	36	3.4	3.4	0
Isoniazid [18, 43]	Lower resistance breakpoint MIC	0.125	0.125	36	3.4	3.4	0
Pyrazinamide [18, 44]	Breakpoint MIC, mg/L	50	50	59	11.2	11.2	0
Moxifloxacin [18, 45, 46]	Breakpoint MIC, mg/L	1	1	16	0.8	0.8	0
% of patients [28, 40]	ADR (Cape Town)	0.68	0.7	142	13.4	13.0	0.4
All (summary)					100	94.43	1.8%

Abbreviations: ADR, acquired drug resistance; AUC, area under the concentration time curve; HFS-TB, hollow fiber system model of tuberculosis; MIC, minimum inhibitory concentration.

Summary of Analyses

HFS-TB Predicted vs. Clinic Observed



Gumbo et al, CID 2015; 61(S1):S25