Pharmacometrics

Assessing the uncertainty and predictive accuracy of mechanistic models; with applications in CML and receptor occupancy

Andy Stein, March 24, 2020
Director of Pharmacometrics

Executive Summary

- **Situation:** Mechanistic and QSP models are more often used to guide drug development and are submitted to the FDA
- **Complication:** QSP models can be extremely complex with many assumptions.
- **Question:** How can these models be validated in a structured way?
- **Solution:** We propose the “Uncertainty Pedigree Table” as a tool for thinking carefully about all sources of uncertainty in the model.

Outline

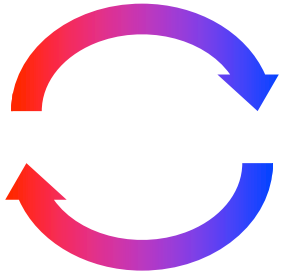
- Definition of QSP
- QSP at Novartis
- Challenges in validating a QSP model
- Introduce credibility assessment and pedigree table
- Two case studies with lessons learned
  - CML case study
  - Receptor Occupancy case study

Definitions of different types of models

	Model Type	Example	Number of differential equations (approximate)	Identifiable parameters. Model can be mathematically analyzed	Can provide mechanistic insights
	Empirical PKPD	Indirect response model	1-5	yes	no
QSP	Mechanistic PKPD	Receptor occupancy, viral kinetics, CML	2-10	yes	yes
QSP	Systems Platform Models	DILIsym, Entelos	11-∞	no	yes

Modeling & Simulation at Novartis

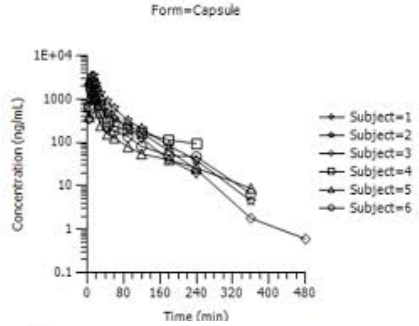
Pharmacokinetic Sciences (PKS)
Modeling & Simulation (M&S)
Includes QSP group
(formed in 2018)

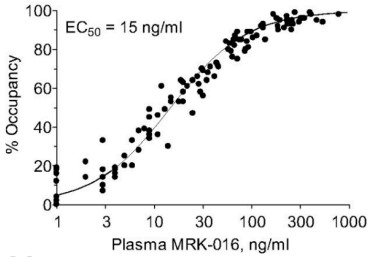


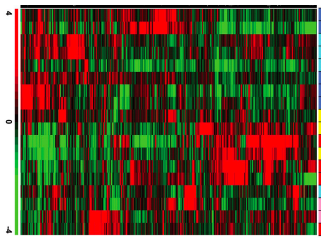
Pharmacometrics (PMX)
QSP use is extremely rare
(formed in 2004)

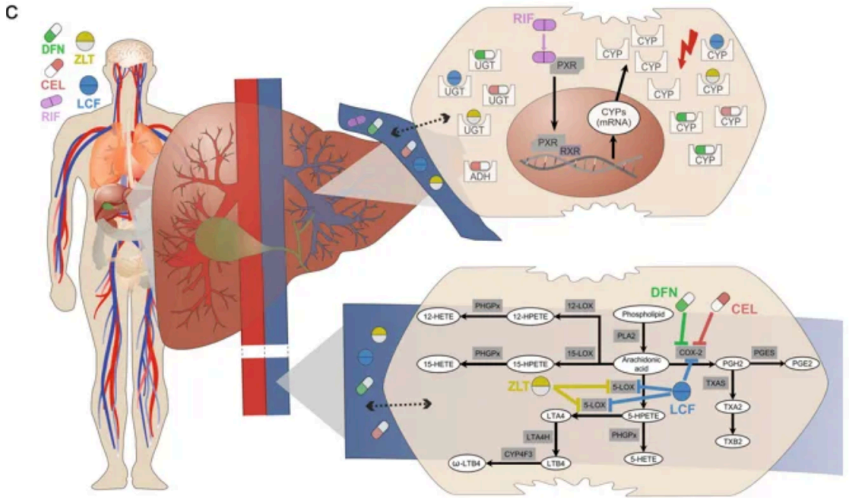
PKS-M&S and PMX are collaborating to
develop strategies for QSP model validation and
to identify opportunities for clinical applications

QSP[1] models integrate data to make predictions









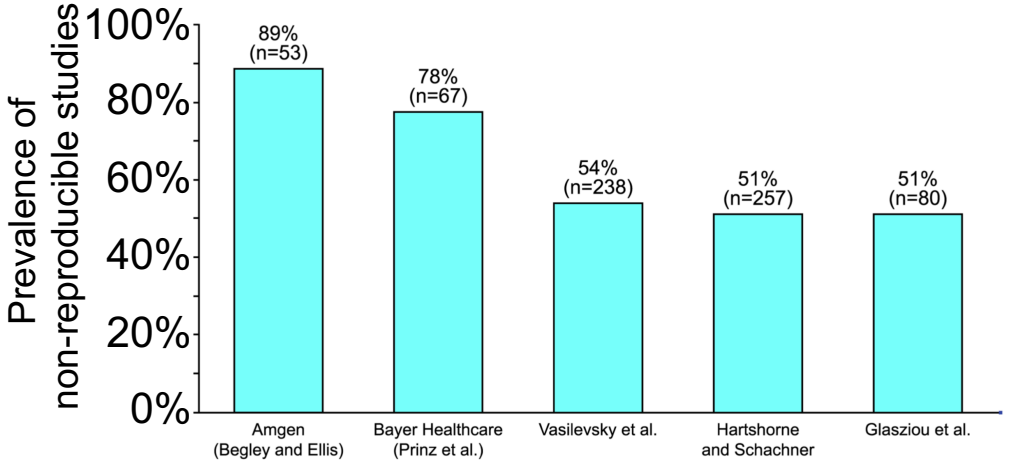
?

But how can we validate that we have the right model?

Reproducibility crisis[1-2]

- Experimental challenges include
  - Cell line misidentification (15-36%)[1]
  - Poor specificity of antibodies
  - Variability in experimental design

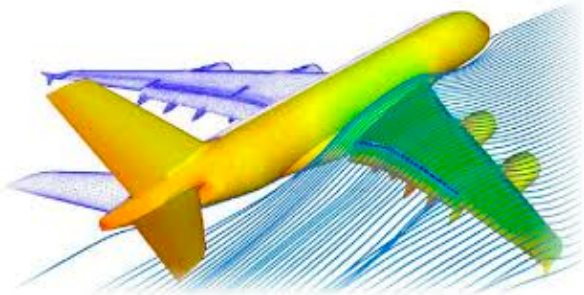
- QSP models often make use of literature data and these experiments were often not validated or reproduced

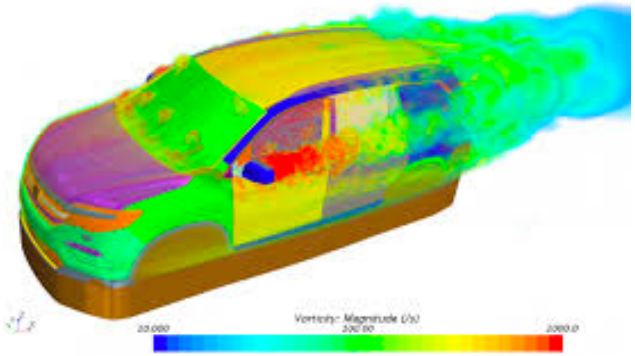


1. Freedman, Leonard P., Iain M. Cockburn, and Timothy S. Simcoe. "The economics of reproducibility in preclinical research." *PLoS Biol* 13.6 (2015): e1002165.
2. Smith, Brian P. "Where Will Statistical Sciences for Clinical Pharmacology Be in 2030?." Clinical Pharmacology & Therapeutics 107.1 (2020): 17-21.

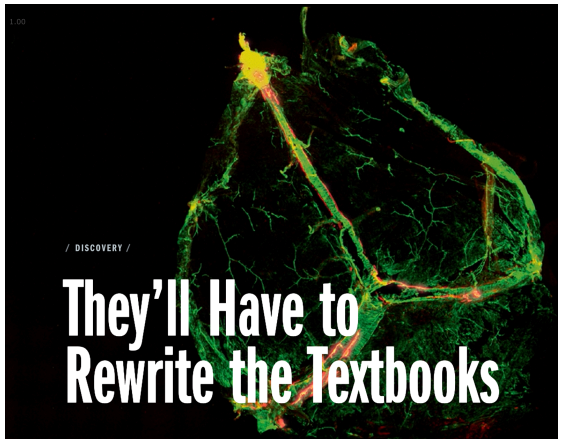
Biological knowledge is rapidly changing

Mechanics/Aerodynamics
(knowledge relatively stable)

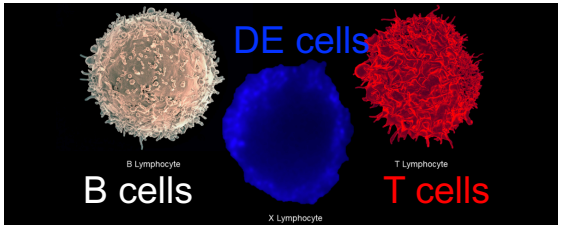




Biology
(knowledge rapidly evolving)



Lymphatics in the brain
https://news.virginia.edu/illimitable/discovery/theyll-have-rewrite-textbooks



DE Lymphocytes
(express BCR and TCR)
https://www.the-scientist.com/news-opinion/novel-type-of-immune-cell-discovered-in-type-1-diabetes-patients-65950

There are no established methods for assessing structural model uncertainty

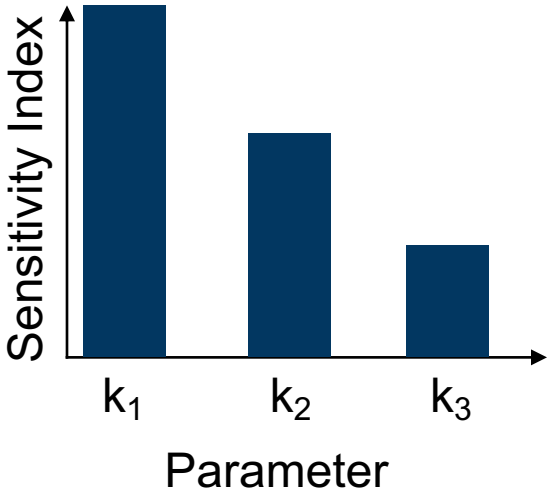
Butterfly Effect	Small changes in initial conditions (or parameters) can give large change in predictions	Sensitivity analysis and virtual patients. Vary parameters and initial conditions over range)

Hawkmoth Effect[1-2]	Small changes in structural model can give large changes in predictions	No systematic method to address structural model uncertainty

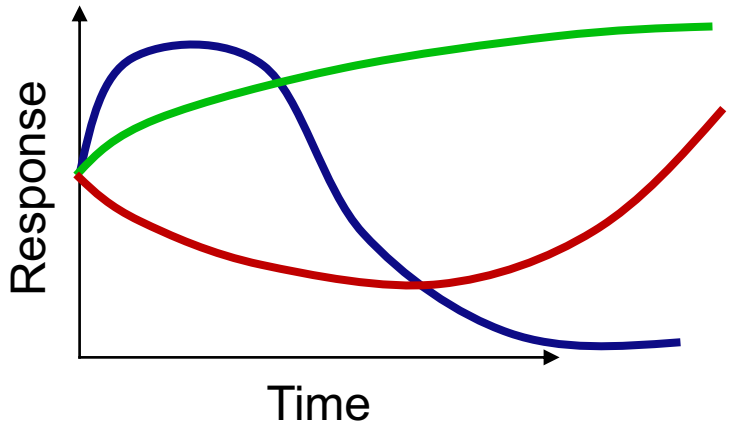
1. Frigg R, et al. “Laplace’s demon and the adventures of his apprentices,” 2014.
2. Thompson E and Smith LA “The hawkmoth effect.” 2014

Sensitivity analysis and virtual patients are standard tools for model evaluation

Global Sensitivity Analysis

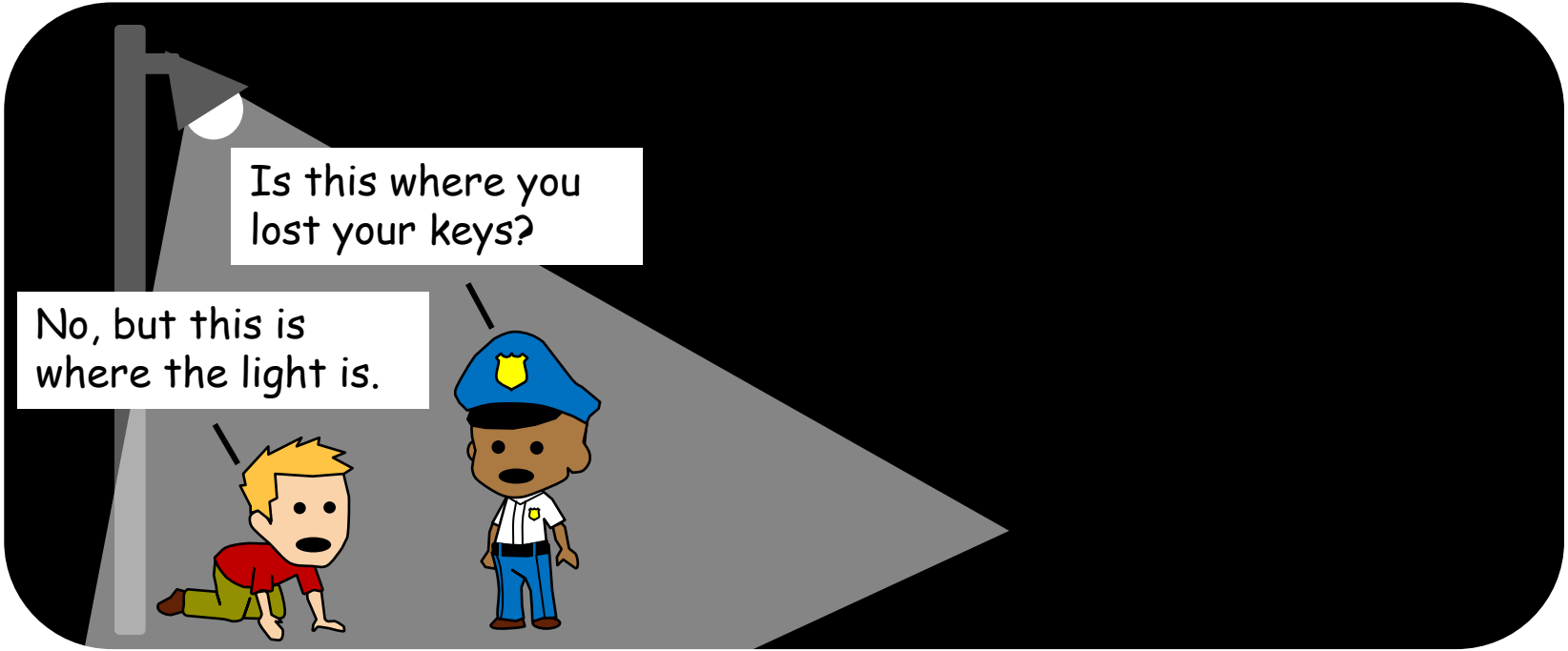


Virtual Patient Simulation

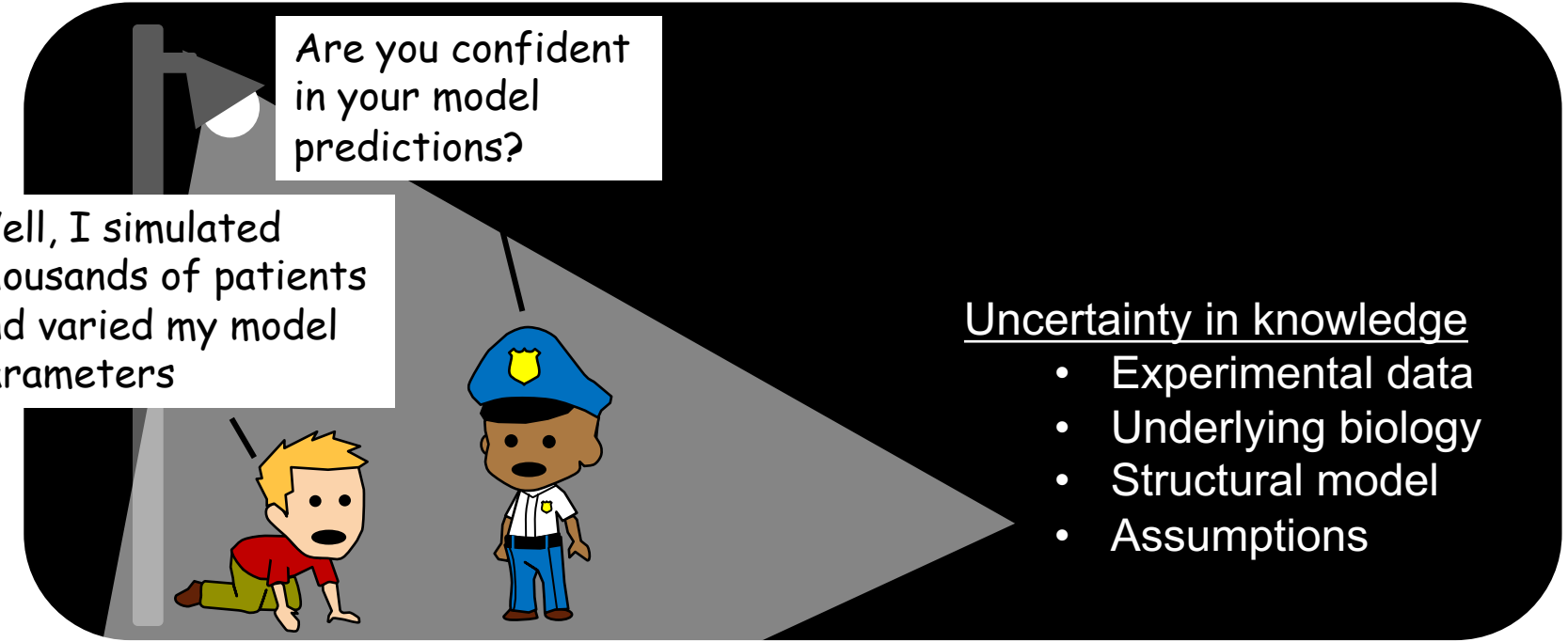


How does the modeling community assess structural model uncertainty?

Streetlight effect



Streetlight effect in model assessment



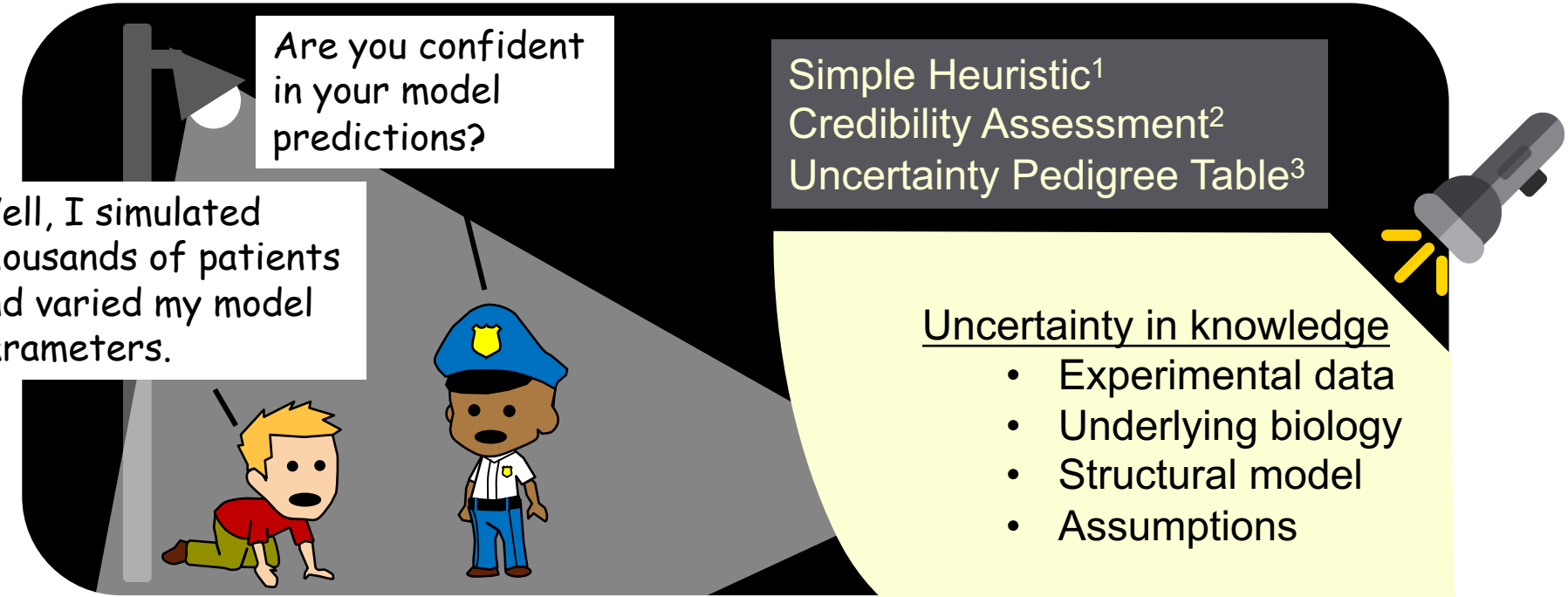
Uncertainty in knowledge

- Experimental data
- Underlying biology
- Structural model
- Assumptions

Uncertainty in parameters

- Sensitivity Analysis
- Virtual Patient Simulation

Streetlight effect in model assessment

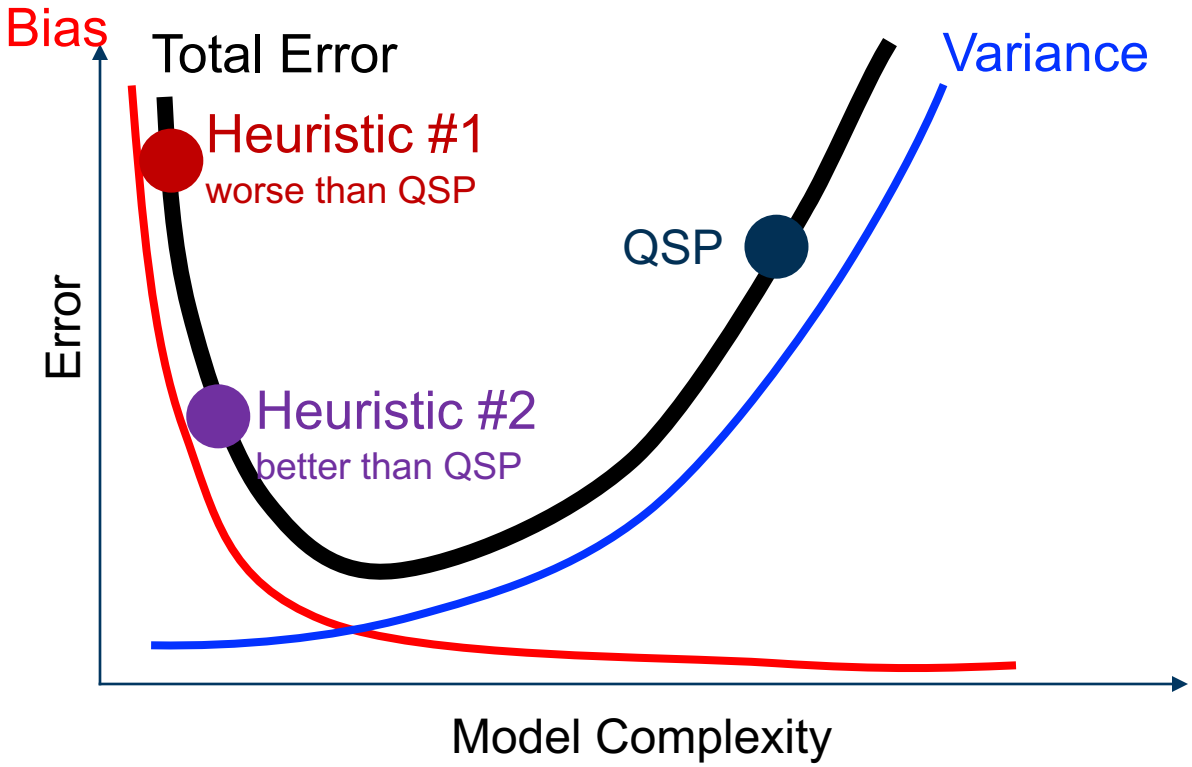


Uncertainty in parameters

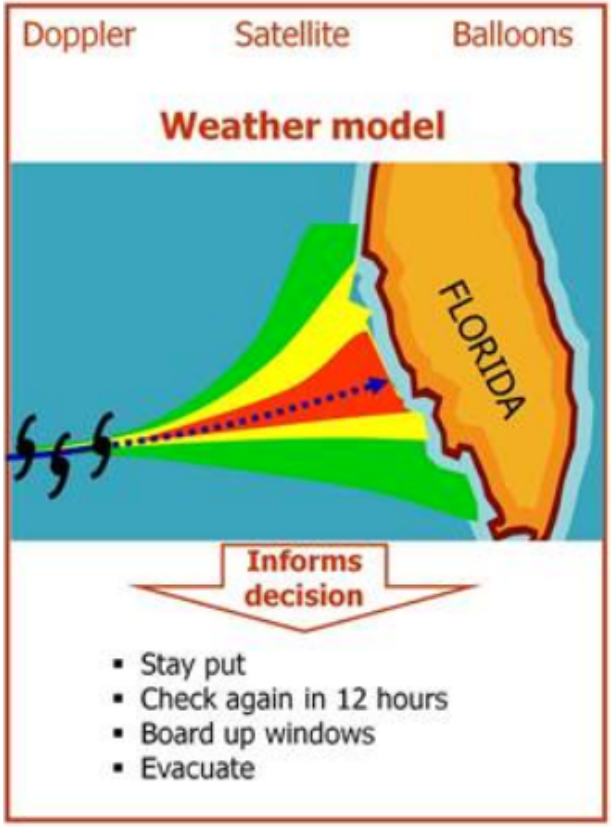
- Sensitivity Analysis
- Virtual Patient Simulation

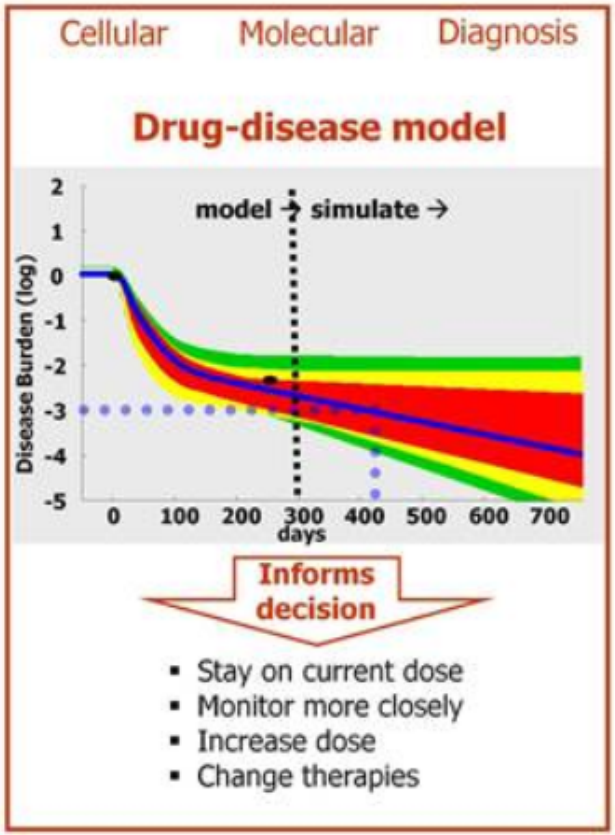
1. Stein and Looby, et al., CPT:PSP, 2018, doi:10.1002/psp4.12311
2. Kuemmel, Colleen, et al. CPT:PSP, 2019 doi:10.1002/psp4.12479.
3. https://opensource.nibr.com/xgx/Resources/Uncertainty_Assessment_Pedigree_Table.pdf

Bias-Variance Tradeoff explains why complex models need careful validation



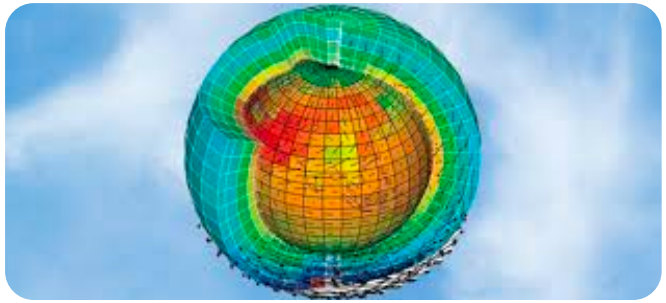
Weather and QSP have many similarities





Bonate, Peter L. *Be a model communicator: and sell your models to anyone*. Peter Bonate, 2014. Figure 29, from Dean Bottino

Daily temperature can be predicted with computational model or historical average

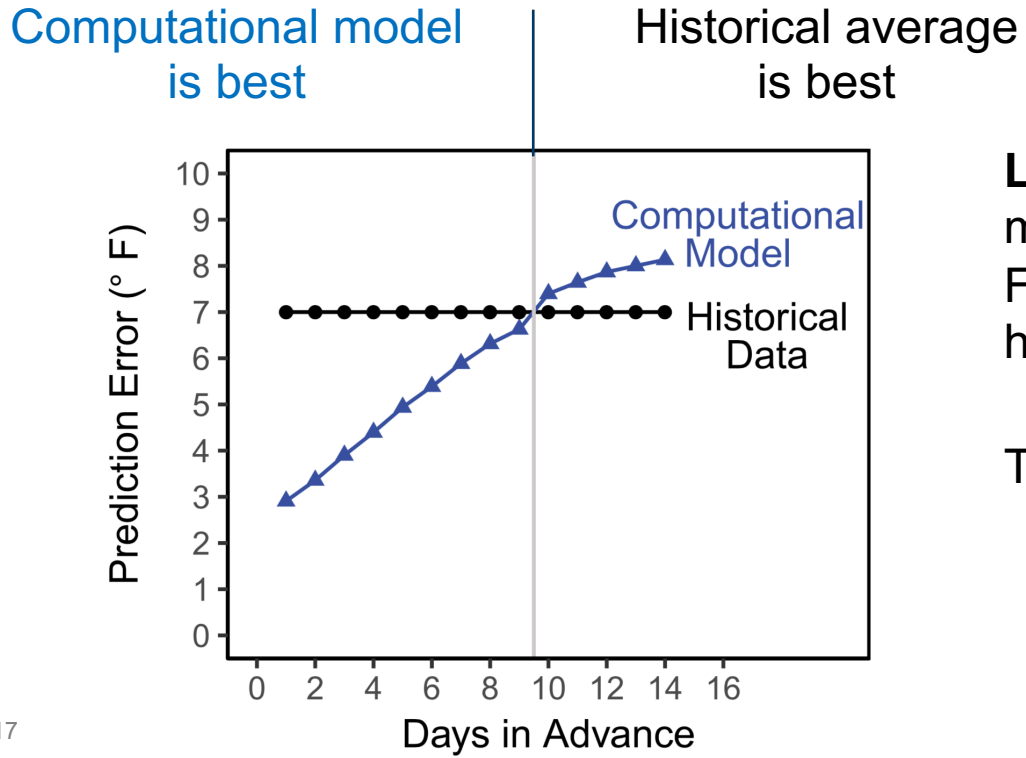


- Computational model (complex)
  - Includes knowledge of atmospheric sciences and large, rich datasets
  - But cannot model everything (e.g. cloud dynamics)
- Historical average (simple)
  - Contains little underlying knowledge of system
  - Easy to implement and understand

Which approach is best?

Silver, Nate. *The signal and the noise: why so many predictions fail--but some don't*. Penguin, 2012.

The best model depends on the type of prediction that is made

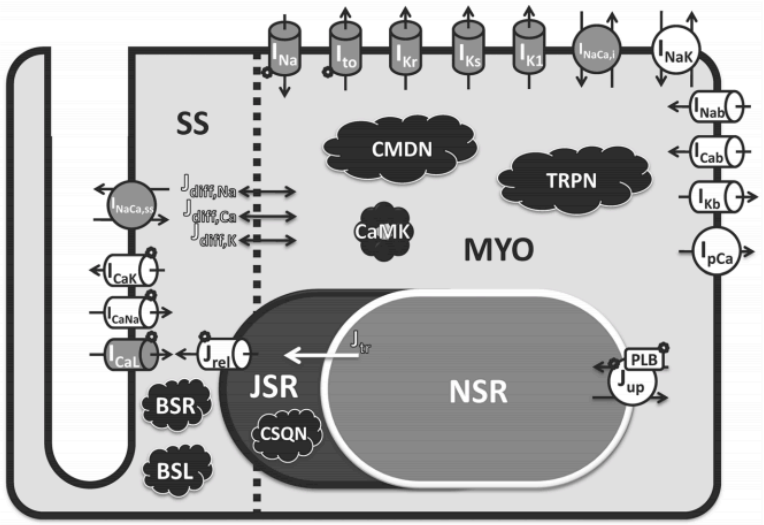


Lesson: For some problems, a mechanistic model performs best. For some problems, a simple heuristic performs best.

There is no way to know in advance

Complex models and simple heuristics exist for predicting cardiotoxicity

"Systems Model"



"Simple Heuristic"

Difference in block between depolarising and repolarising ion-channels

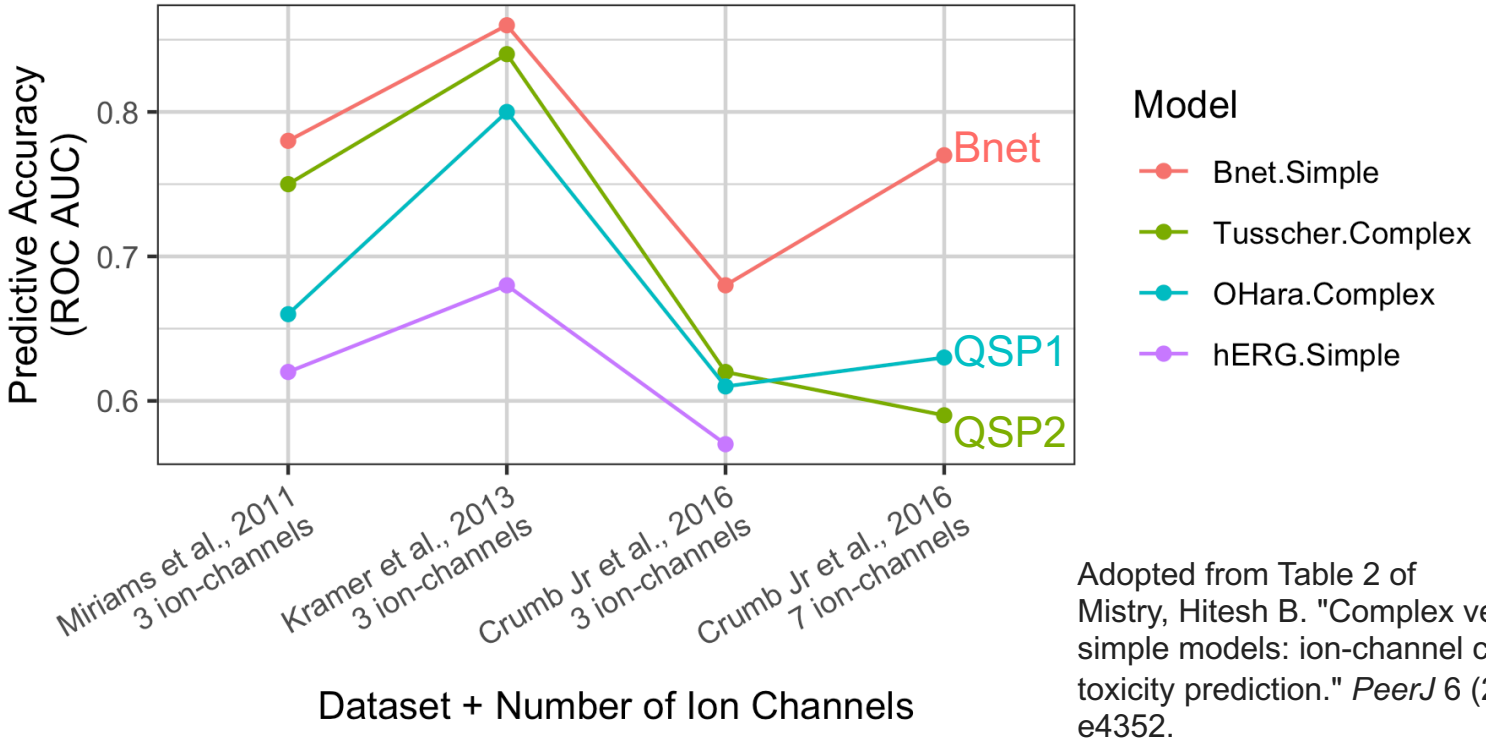
repolarizing depolarizing

$$B_{net} = \sum_{i=1}^{n} R_i - \sum_{j=1}^{m} D_j$$

All regression coefficients = 1
i.e. all have equal importance

Mistry, Hitesh B. "Complex versus simple models: ion-channel cardiac toxicity prediction." *PeerJ* 6 (2018): e4352.

Simple heuristics and complex biophysical models have similar predictive accuracy



Adopted from Table 2 of Mistry, Hitesh B. "Complex versus simple models: ion-channel cardiac toxicity prediction." *PeerJ* 6 (2018): e4352.

Simple heuristics are available for some questions[1]

Item to predict	Simple Heuristic	Complex Model
Exposure in humans	Allometric Scaling Wajima Superposition[2]	Full PBPK
Cardiotoxicity	Bnet[3]	Systems model
Benefit of oncology drug combination	Assume independence[4]	QSP model
Predicting drug induced liver injury	?	DILIsym

1. Stein A and Looby M, et al., CPT:PSP, 2018, doi:10.1002/psp4.12311
2. Lombardo, F, et al. J Pharmaceutical Sci 105.3 (2016): 1277-1287.
3. Mistry, HB. PeerJ 6 (2018): e4352.
4. Palmer, AC and Sorger PK. Cell 171.7 (2017): 1678-1691.

Proposal: when simple heuristics are not available, we have developed the Uncertainty Pedigree Table

Frameworks exist for assessing model uncertainty.

WHITE PAPER

Consideration of a Credibility Assessment Framework in Model-Informed Drug Development: Potential Application to Physiologically-Based Pharmacokinetic Modeling and Simulation

Colleen Kuemmel[1,*], Yuching Yang[1], Xinyuan Zhang[1], Jeffry Florian[1], Hao Zhu[1], Million Tegenge[2], Shiew-Mei Huang[1], Yaning Wang[1], Tina Morrison[3] and Issam Zineh[1]

Cheat sheet for model uncertainty assessment Version 1.01

Novartis Contributors: Andrew M Stein, Jeff D Kearns, Jaeyeon Kim, Alison Margolskee

Context of use table provides guide for how carefully model should be validated

Question	The specific question, decision, or concern that is being addressed
Context of Use	Specific role model plays in addressing question
Model Influence	Contribution model makes to decision (low, medium, high)
Model Risk	Likelihood model lead to incorrect decision (low, medium, high)
Decision Consequence	Severity of the adverse outcome from incorrect decision
Validation Plan	We propose to use the uncertainty pedigree table

Kuemmel, Colleen, et al. "Consideration of a Credibility Assessment Framework in Model-Informed Drug Development: Potential Application to Physiologically-Based Pharmacokinetic Modeling and Simulation." CPT: pharmacometrics & systems pharmacology (2019).

Pedigree table provides guide for assessing model uncertainty

Used previously for:

- Climate change
- Health impacts of incinerator
- Ground water protection from pollution

Combining Quantitative and Qualitative Measures of Uncertainty in Model-Based Environmental Assessment: The NUSAP System

Jeroen P. van der Sluijs,[1]* Matthieu Craye,[2] Silvio Funtowicz,[2] Penny Kloprogge,[1] Jerry Ravetz,[3] and James Risbey[4]

A framework for dealing with uncertainty due to model structure error

Jens Christian Refsgaard [a,*], Jeroen P. van der Sluijs [b], James Brown [c], Peter van der Keur [a]

Pedigree table idea: provide qualitative assessment for each source of uncertainty

1. Experimental data
2. Consensus on biology
3. Structural model exploration
4. Assumption accuracy
5. Interpolation
6. Extrapolation
7. Inter-subject variability
8. Parameter sensitivity
9. Quality control

Each criteria is assigned a qualitative score of 0-4

Recommendation: scores assessed during scoping and again after modeling is completed

https://opensource.nibr.com/xgx/Resources/Uncertainty_Assessment_Pedigree_Table.pdf

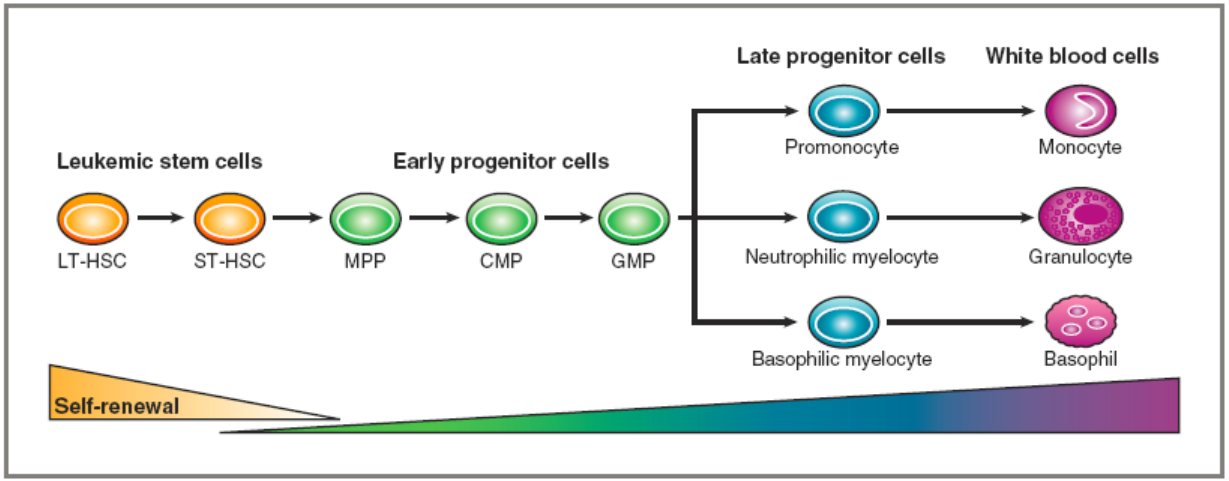
Example: Experimental data quality

Criteria	Score	Assessment
Very Good	4	All data from validated assays and reproducible experiments. Data fully describes system
Good	3	Enough trusted data for good description of system
Fair	2	Limited data with old experiments using unvalidated assays
Poor	1	Very limited data, educated guesses inform the model
Very Poor	0	No data

Uncertainty pedigree table is easy to use

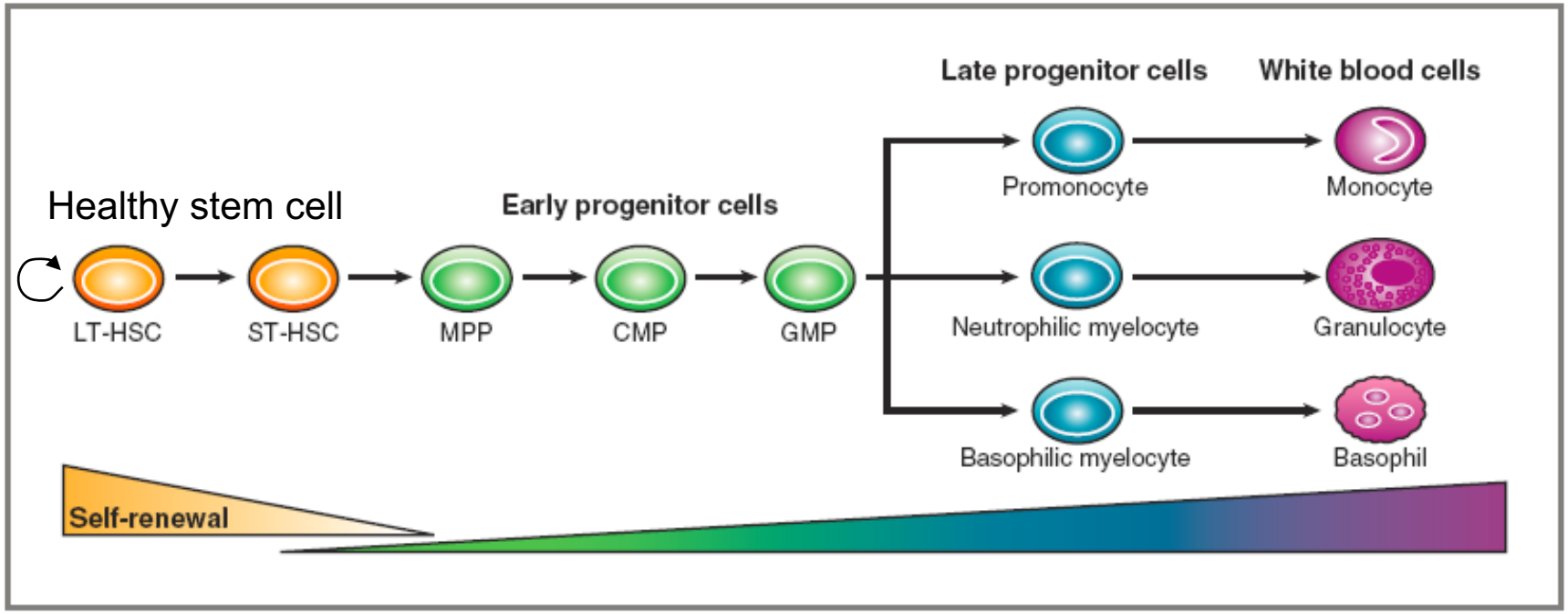
- The pedigree table can be completed at a high level, or a detailed level (for each dataset, parameter, and assumption)

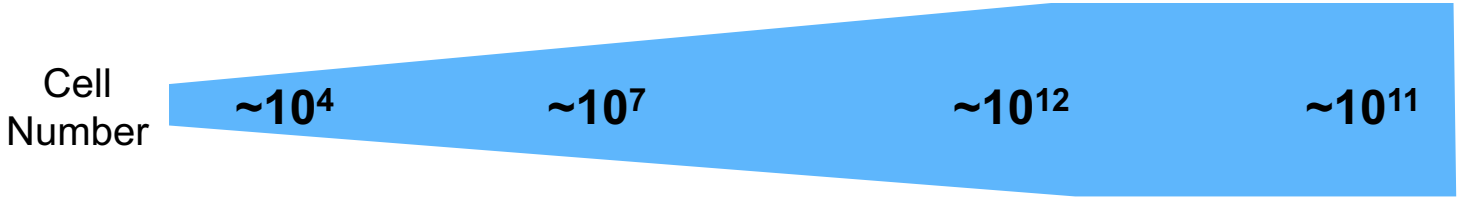
- Table is easy to generate, understand, and discuss with colleagues of different backgrounds



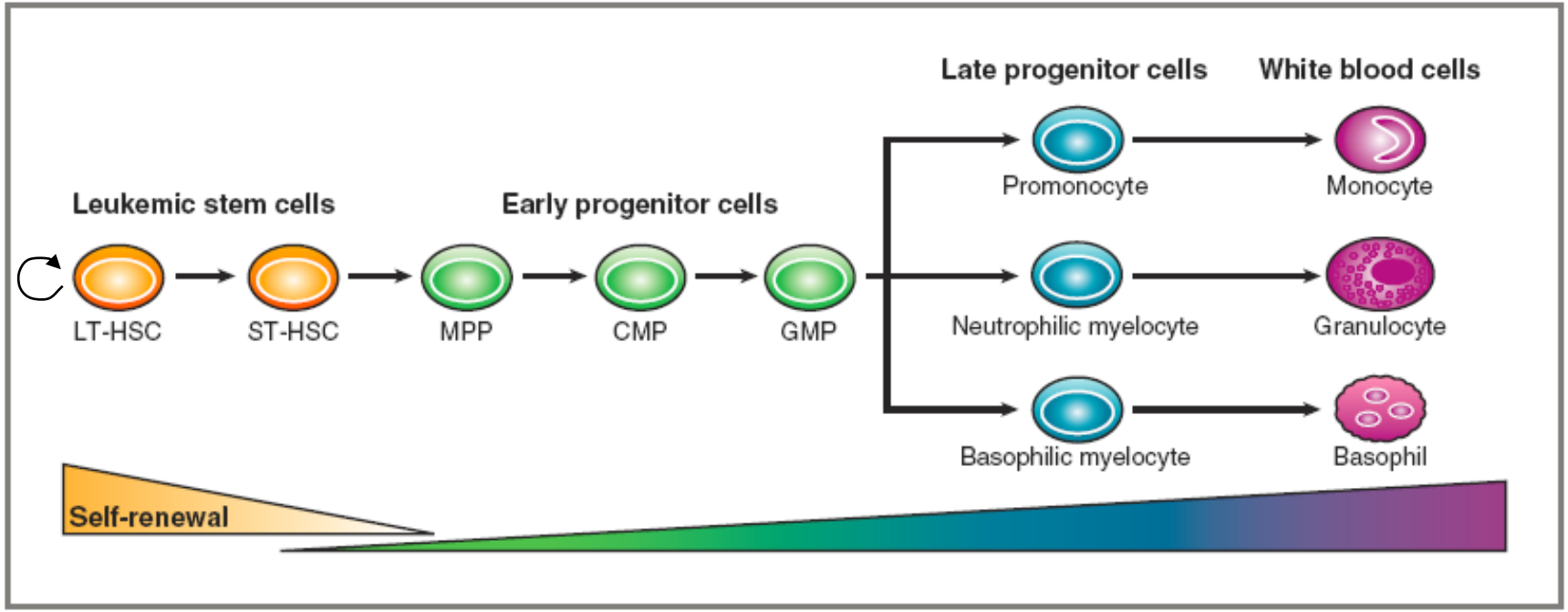
Predicting whether Gleevec therapy can eliminate leukemic stem cells and safely be stopped

Hematopoiesis





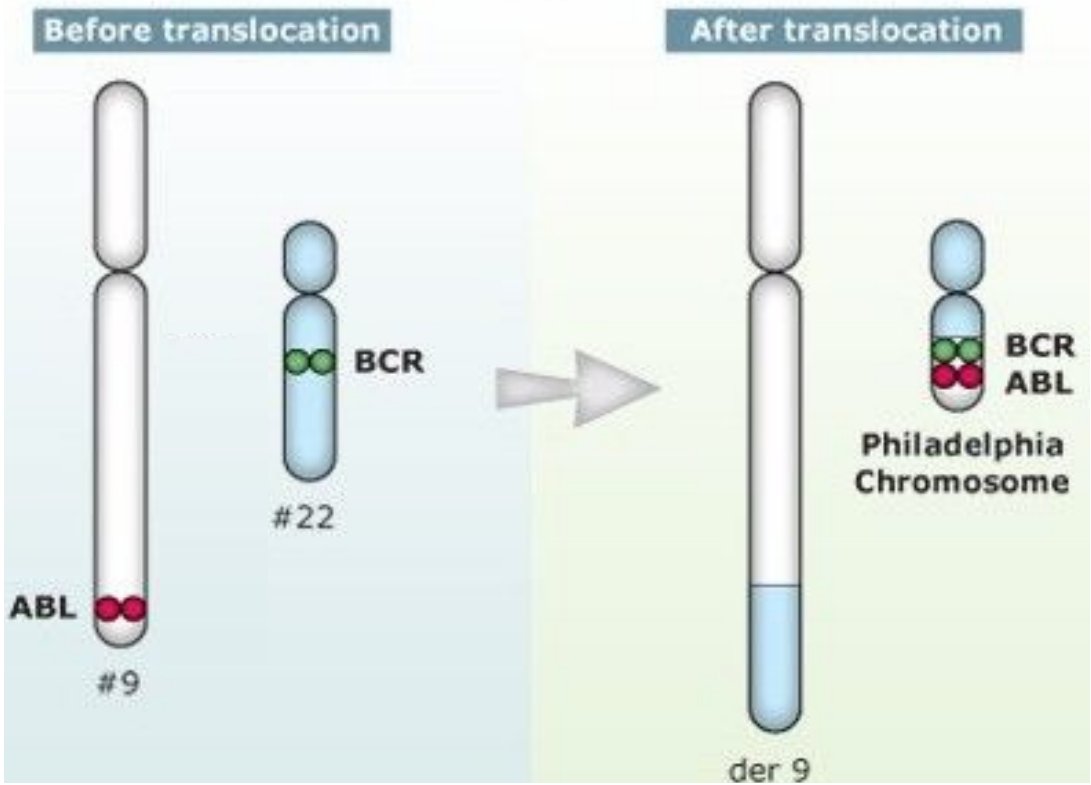
Leukopoiesis in Chronic Myeloid Leuk.



Cell Number

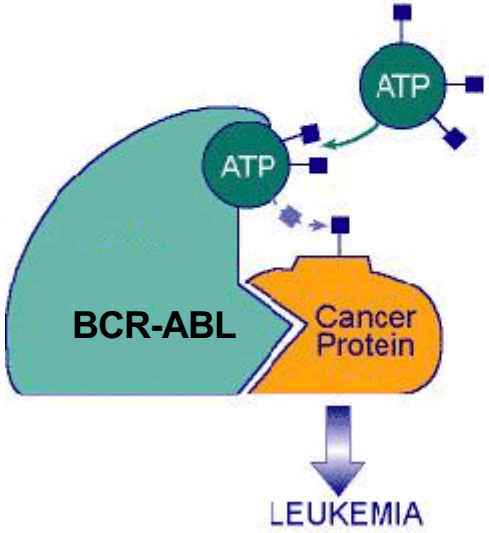
~10^4 ~$4 \cdot 10^7$ ~$10 \cdot 10^{12}$ ~$10 \cdot 10^{11}$

BCR-ABL – Translocation

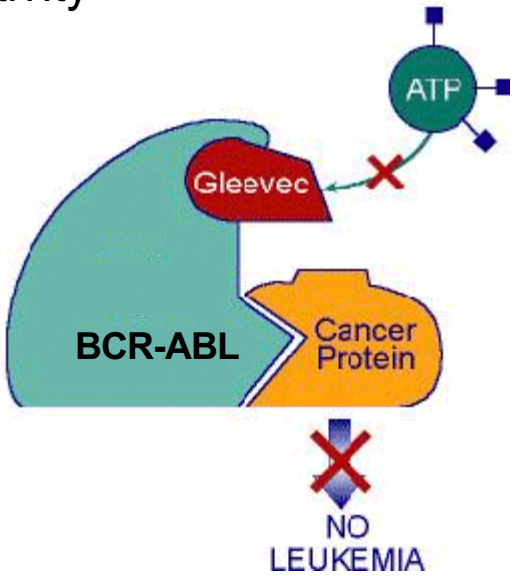


Function of BCR-ABL and Gleevec

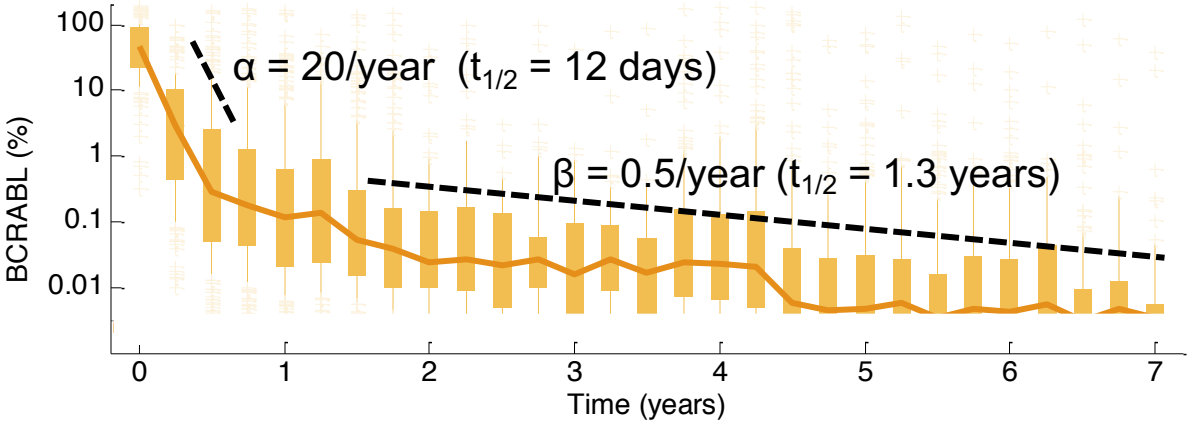
BCR-ABL:
constitutively active kinase that triggers cell proliferation

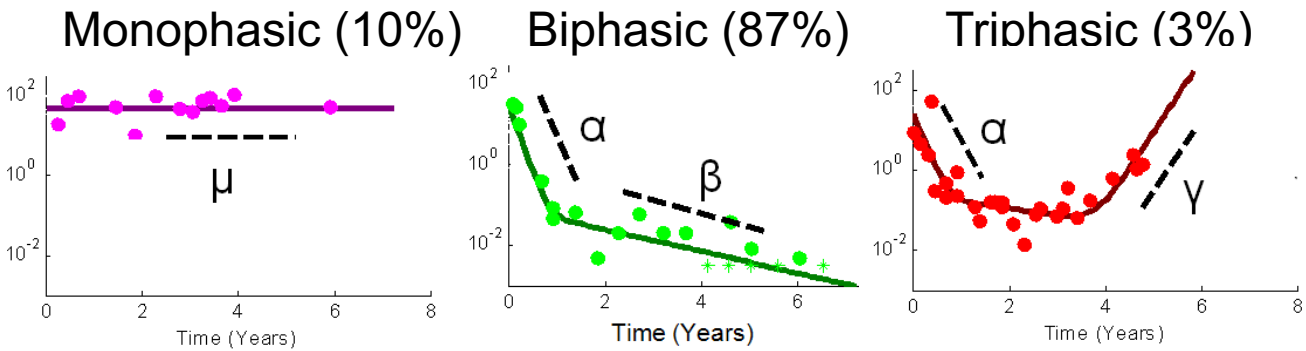


Gleevec:
binds ATP binding site on BCR-ABL, inhibiting its activity

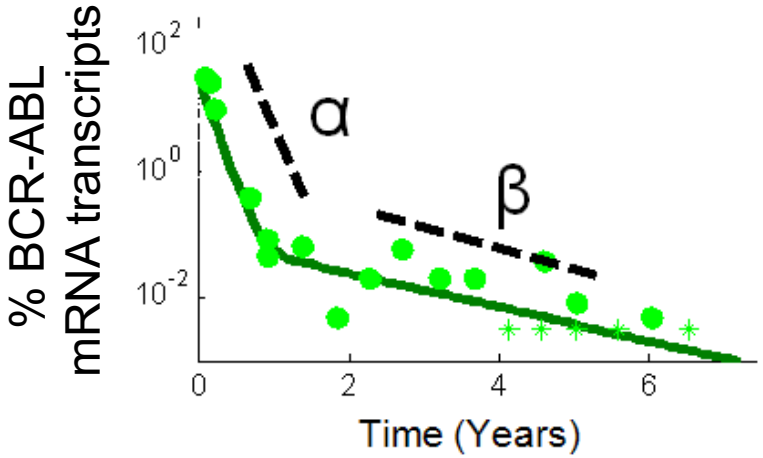


During Gleevec Therapy, patients exhibit a biphasic decline in BCR-ABL mRNA transcripts





Does Gleevec impact leukemic stem cells (LSCs)?



Elimination of stem cells is needed for sustained response after Gleevec discontinuation

History of Gleevec Modeling

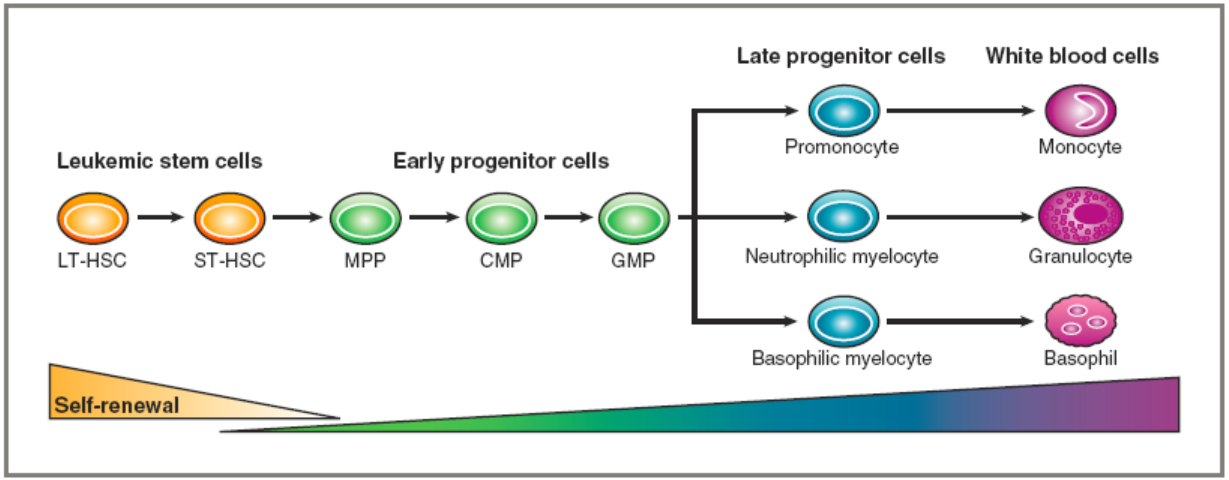
- 2001: Gleevec approved
- 2005: Gleevec modeling began at Novartis.
- 2005: First Gleevec modeling paper appeared in Nature[1].
  Predicted Gleevec did not impact stem cells.
- 2006: Second Gleevec modeling paper appeared in Nature Mecidine[2].
  Predicted Gleevec did impact stem cells.
- 2007: Investigator initiated Gleevec discontinuation trial.
- 2009: Initial results of discontinuation trial. ~40% maintained response.
- 2009-2011: Gleevec modeling presentation[3] and paper[4] evaluating multiple models, providing further supporting stem cell effect

1. Michor et al. Nature 2005; 435:1267-1270.
2. Roeder et al., Nature Medicine 2006; 12:1181-4
3. Bottino et al., Blood 2009; 114, 209, Abs #506
4. Stein et al., Clinical Cancer Research 2011; 17:6812-21

Context of use for Gleevec/CML model

Question	Will a patient ever be able to discontinue Gleevec therapy without relapse?
Context of Use	Mathematical was fit to data and simulations used to predict what is happening to the leukemic stem cells.
Model Influence	**Low/None.** Investigators decided to run a carefully monitored discontinuation study
Model Risk	**None.** Because model did not (to our knowledge) influence the investigator's decision to initiate study, there was no risk of the model leading to an incorrect decision.
Decision Consequence	**Low.** Careful monitoring of patients that discontinue reduces the overall impact on the patients.
Validation Plan	It is instructive to apply the Pedigree Table retrospectively.

Model-based description of hematopoiesis



LSC-P: proliferating stem cell

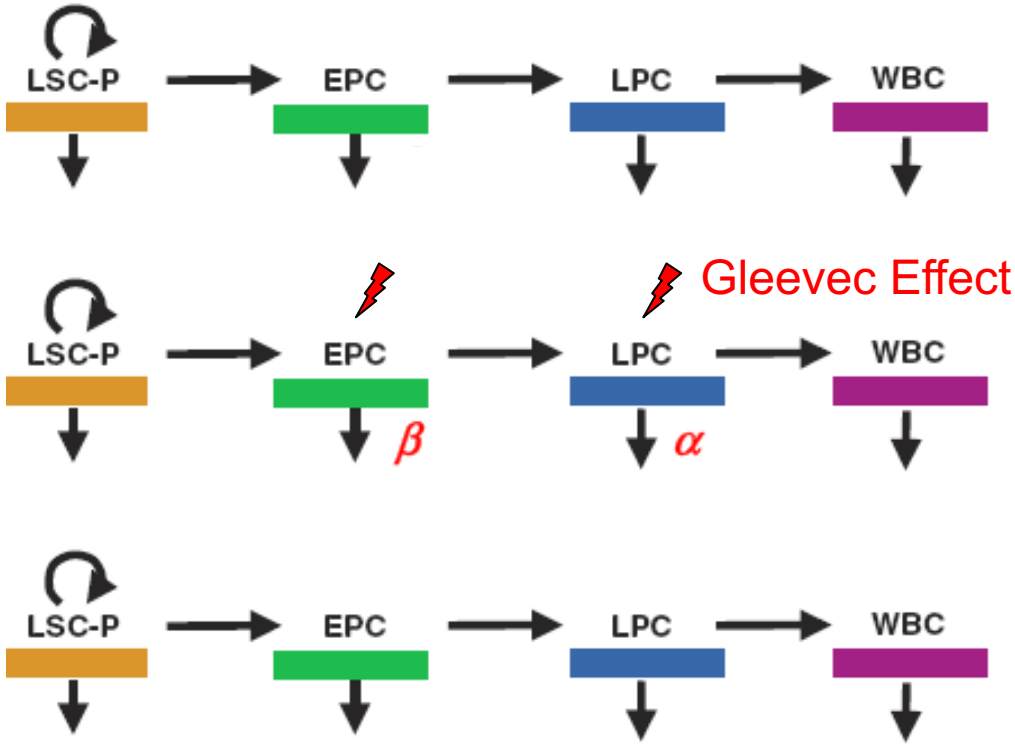
EPC: early progenitor cell
LPC: late progenitor cell
WBC: white blood cell

First mechanistic model of gleevec assumed gleevec had no impact on stem cells (2005)

Healthy Cells (H)

Leukemic Cells (L)

Leukemic Cells with resistance (R)



1. Michor et al. Nature 2005; 435:1267-1270.

Measurement model for linking WBC populations to BCR-ABL assay measurement

Healthy Cells (H) — 2 ABL

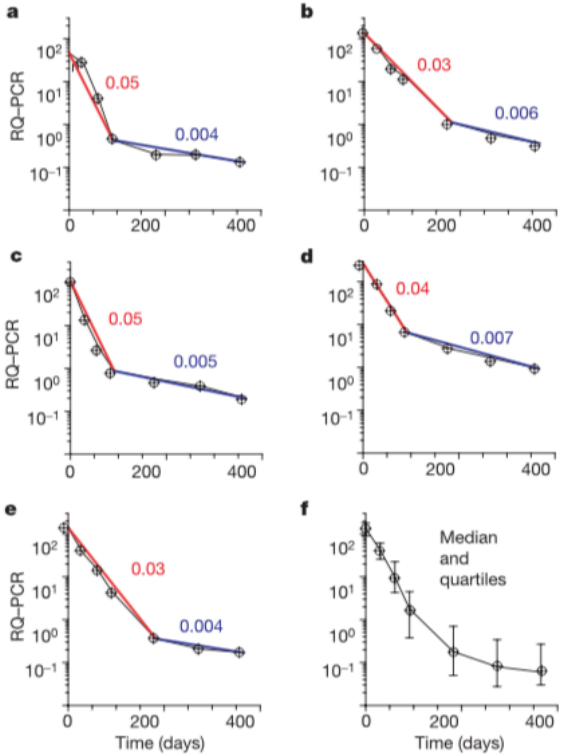
Leukemic Cells (L) — 1 ABL, 1 BCR-ABL

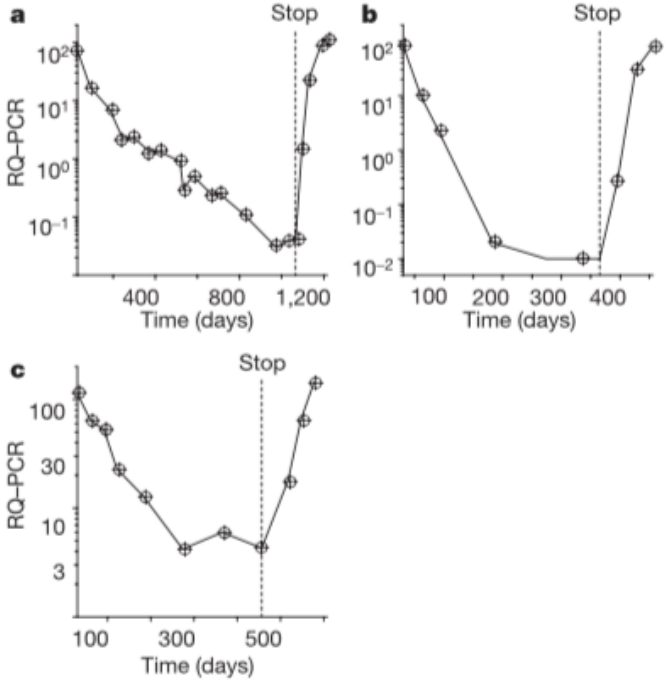
Leukemic Cells with resistance (R) — 1 ABL, 1 BCR-ABL

$$BCR-ABL\% = \frac{BCR-ABL\ mRNA}{ABL\ mRNA}$$

$$= \frac{L+R}{2H+L+R}$$

Model fits with no impact on stem cells were consistent with available data[1]



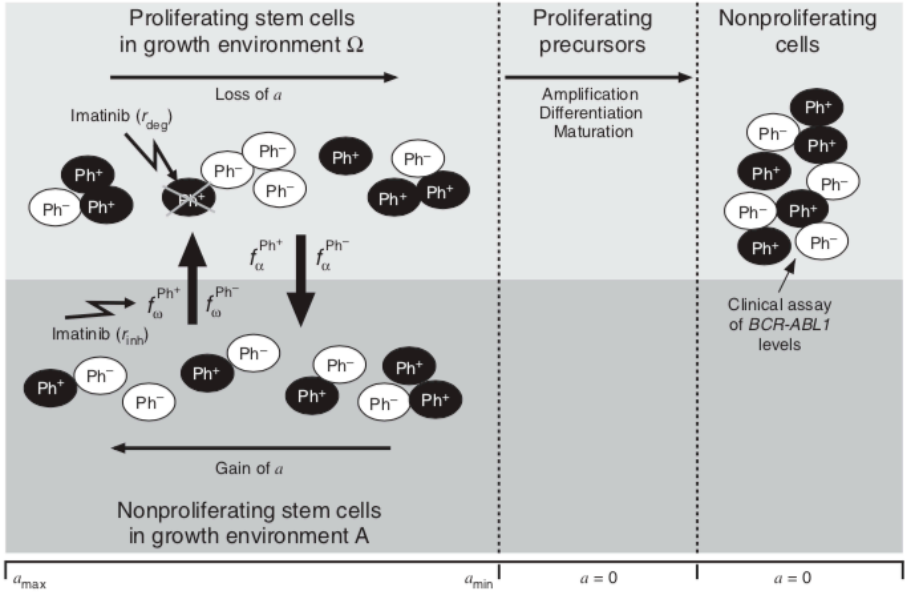


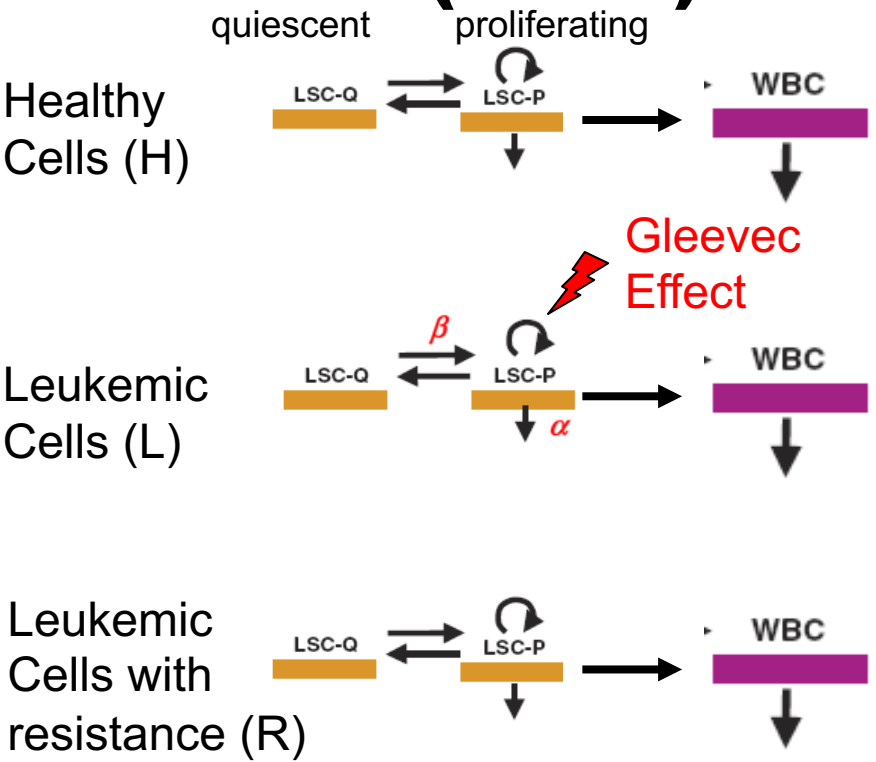
1. Michor et al. Nature 2005; 435:1267-1270.

Pedigree Assessment of Model

Criteria	Score	Description
Data	4 Very Good	BCR-ABL assay was run by excellent lab with data from many patients.
Biological Understanding	1 Poor	Embryonic field. There was consensus on the overall hematopoietic process. But the impact of Gleevec on hematopoiesis was not established.
Structural Model Exploration	0 Very Poor	Only one model was considered.
Assumptions	1 Poor	One key assumption was that Gleevec did not impact stem cells. The authors never explored relaxing this assumption.

Second model of Gleevec assumed gleevec impacted stem cells[1] (2006)





1. Roeder et al., Nature Medicine 2006; 12:1181-4

Measurement model for linking WBC populations to BCR-ABL assay measurement [same as previous]

Healthy Cells (H) — 2 ABL

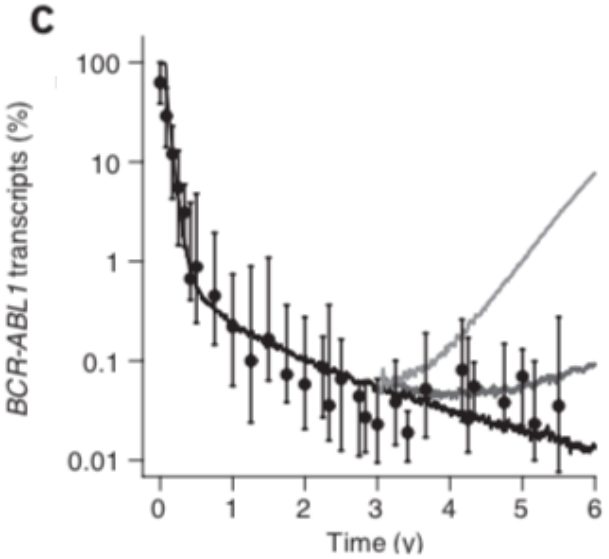
Leukemic Cells (L) — 1 ABL, 1 BCR-ABL

Leukemic Cells with resistance (R) — 1 ABL, 1 BCR-ABL

$$BCR-ABL\% = \frac{BCR-ABL\ mRNA}{ABL\ mRNA}$$

$$= \frac{L+R}{2H+L+R}$$

Model fits with gleevec impact on stem cells were consistent with available data[1]



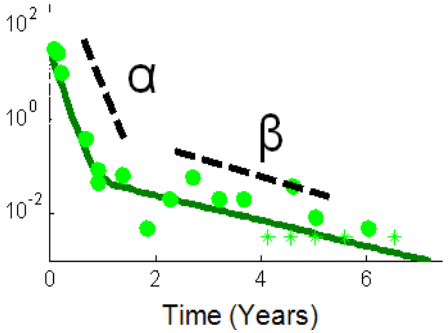
1. Roeder et al., Nature Medicine 2006; 12:1181-4

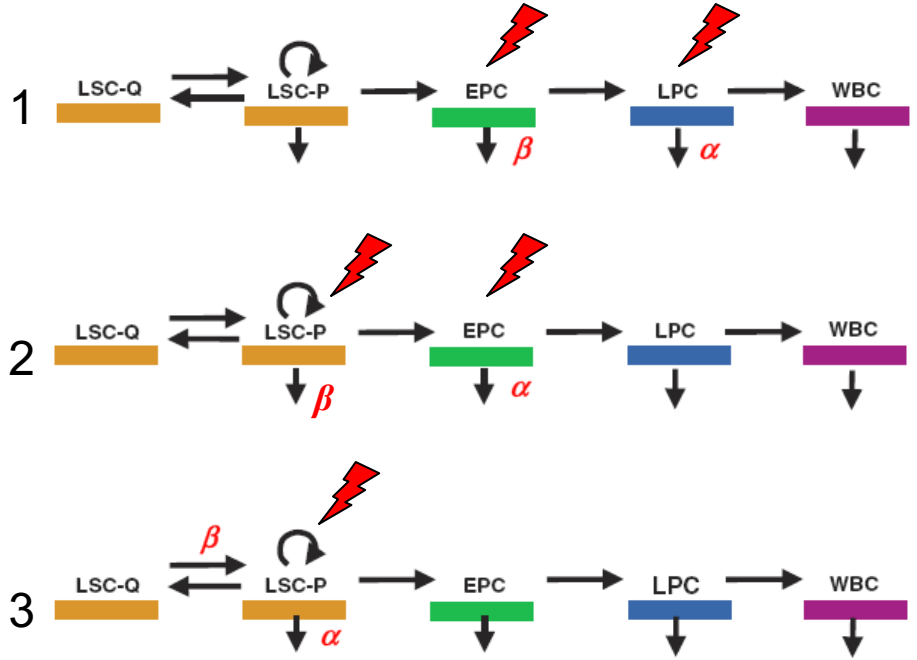
Pedigree Assessment of Model

Criteria	Score	Description
Data	4 Very Good	BCR-ABL assay was run by excellent lab with data from many patients.
Biological Understanding	1 Poor	Embryonic field. There was consensus on the overall hematopoietic process. But the impact of Gleevec on hematopoiesis was not established.
Structural Model Exploration	0 Very Poor	Only one model was considered in paper.
Assumptions	1 Poor	Key assumption was Gleevec impact stem cells. The authors never explored relaxing this assumption.

Three different hypothesis describe the biphasic decline in BCR-ABL[0]

LSC-Q: quiescent leuk. stem cell
LSC-P: proliferating leuk. stem cell
EPC: early progenitor cell
LPC: late progenitor cell
WBC: white blood cell

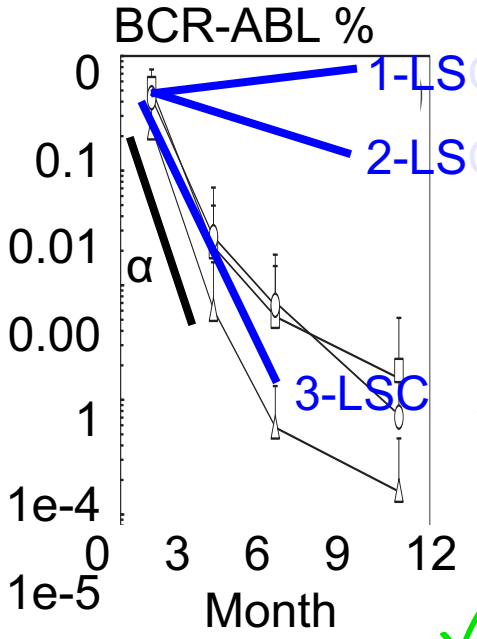


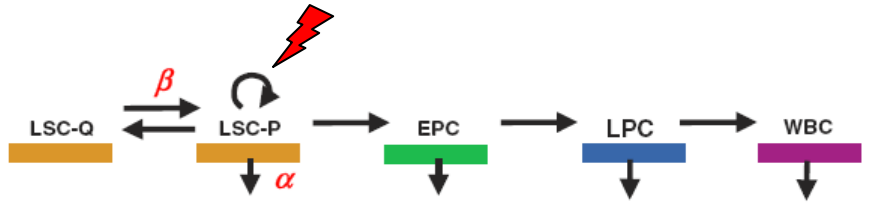


0. Stein et al., Clinical Cancer Research 2011; 17:6812-21

1. Michor et al, Nature, 435, 1267 (2005)
2. Bottino et al, J. Clin. Oncol, 27, 7056 (2009)
3. Roeder et al., Nature Medicine, 12, 1181 (2006)

Additional measurements in the bone marrow[1-2] help identify correct model





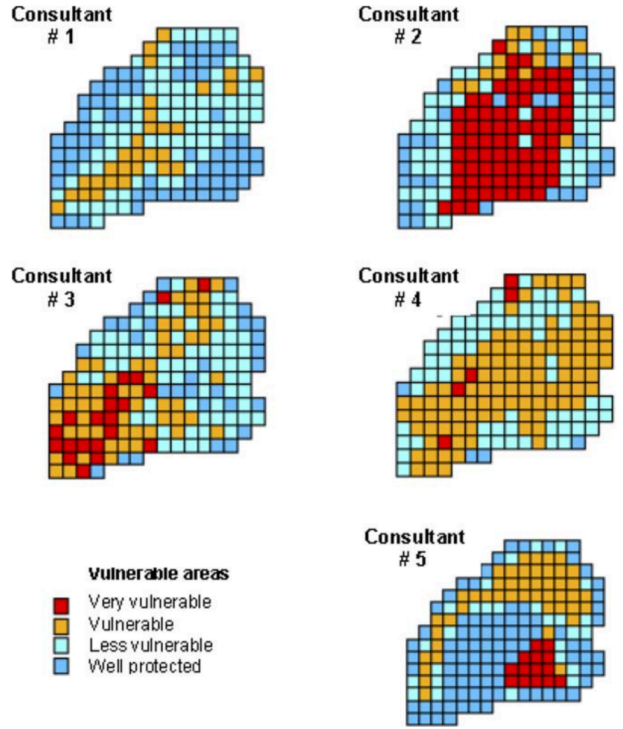
1. Abe et al., *Int J. Hematol*, 88, 471 (2008)
2. Minami et al., *Leukemia,* 26, 2142 (2012)

Pedigree Assessment of Model

Criteria	Score	Description
Data	4 Very Good	BCR-ABL assay was run by excellent lab with data from many patients.
Biological Understanding	2 Fair	Additional evidence now available for Gleevec's impact on stem cells
Structural Model Exploration	2 Fair	Three structural models were considered. But what about other processes?
Assumptions	2 Fair	Gleevec affect on stem cells was explored. But what about other assumptions (Model had no feedback and no interaction with immune system. was this ok?)

How to explore multiple structural models?

- Idea: have multiple independent modelers.
- For most PopPK datasets, it's likely that multiple independent modelers would come up with similar predictions.
- For a QSP system, how do predictions vary across modelers?
  - We don't know.
  - We almost never look.



Model predictions on aquifer vulnerability towards nitrate pollution in area west of Copenhagen. Predictions from 5 different consultants[1]

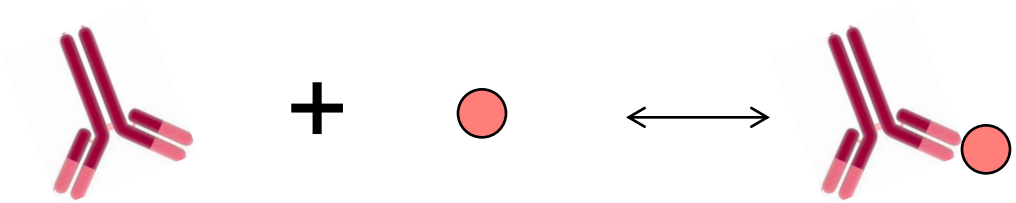
1. Refsgaard et al., Advances in Water Resources, 29, 1586, 2006

Summary of Uncertainty Assessment

- At first, multiple hypotheses were all consistent with the available data.
- Considering multiple structural models was essential for evaluating model uncertainty.
- The Michor et al. and Roeder et al. papers were very valuable scientific contributions.
- However, these contributions were not designed to carefully inform a decision.
- Since these models did not impact a decision, low scores in the uncertainty assessment were acceptable.

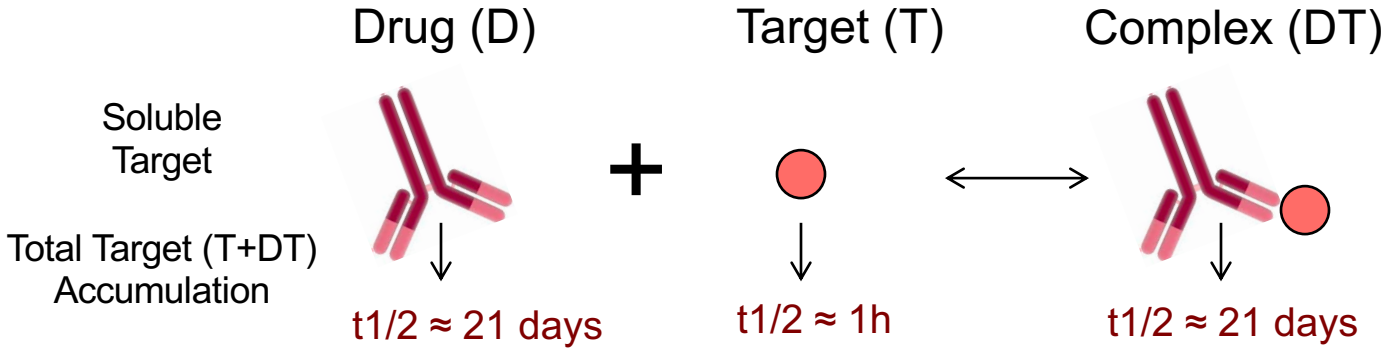
CML model is simpler than traditional QSP Same issue arises with QSP

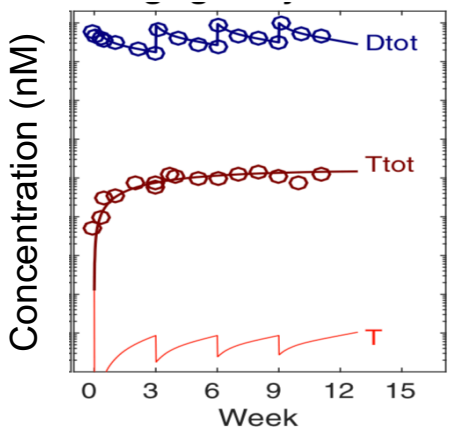
- The case studies presented here were simpler than the typical QSP model
- But the issues around uncertainty of the model structure and the assumptions apply there as well.
- Uncertainty Pedigree Table is useful for having a productive discussion with all stakeholders.



Predicting receptor occupancy to guide dose selection

Total soluble target accumulation





AS

Phase 1 all-comers immuno-onc trials may not provide enough data for dose selection

- **No observed dose-safety relationship:**
  Grade 3/4 adverse events are rare
- **No observed dose-efficacy relationship:**
  Phase 1 “all comers” trial
- **No established biomarkers for response**
- Therefore, a prediction of target engagement was used to guide dose selection for Phase 2 expansion.

Atezolizumab used PK and biodistribution to predict receptor occupancy in the tumor

2.2.2 What is the basis of the dose selection?

The serum atezolizumab concentration of 6 µg/mL was set as a target serum concentration based on the nonclinical tissue distribution data in tumor-bearing mice and receptor occupancy in the tumor.

In pivotal Phase 2 trial used the fixed 1200 mg dosage q3w, steady-state trough serum concentrations for all patients were above the targeted serum concentration of 6 µg/mL.

From the FDA Clinical Pharmacology Atezolizumab Review

Atezolizumab predicted Phase 2 dose with a simple mathematical model + preclinical tumor data[1].

- Atezolizumab drug concentration ($C_{trough,ss}$) data from clinical trial.
- Binding affinity (Kss) was estimated from a cell based assay
- Biodistribution coefficient to tissue is assumed to be 30% (B = 0.3) based on preclinical tumor model
- Avastin combo thought to decrease tumor penetration by additional 30% (factor of A = 0.7).

$$\text{Target Engagement} = \frac{C_{tumor,ss}}{K_{ss} + C_{tumor,ss}} = \frac{A \cdot B \cdot C_{trough,ss}}{K_{ss} + A \cdot B \cdot C_{tumor,ss}}$$

1. Deng, Rong, et al. "Preclinical pharmacokinetics, pharmacodynamics, tissue distribution, and tumor penetration of anti-PD-L1 monoclonal antibody, an immune checkpoint inhibitor." *MAbs*. Vol. 8. No. 3. Taylor & Francis, 2016, Suppl Fig 5

Context of use for receptor occupancy guided dose selection

Question	What dose should be tested in a Pivotal Phase 2 study? Is it sufficient to test one dose?
Context of Use	Modeling used to predict dose that gave over 95% receptor occupancy
Model Influence	Medium. Safety data is available. Model helpful in determining when to stop exploration.
Model Risk	Medium. If model assumptions are incorrect, it could lead to a suboptimal dose selection
Decision Consequence	A suboptimal dose could be selected
Validation Plan	Retrospectively, apply the Uncertainty Pedigree Table

Kuemmel, Colleen, et al. "Consideration of a Credibility Assessment Framework in Model-Informed Drug Development: Potential Application to Physiologically-Based Pharmacokinetic Modeling and Simulation." CPT: pharmacometrics & systems pharmacology (2019).

Pedigree Assessment of Model

Criteria	Score	Description
Data	2 Fair	PK data is robust, Limited data on drug biodistribution to tumors in clinic
Biological Understanding	3 Good	General consensus on receptor binding kinetics, but receptor turnover and shedding was not addressed
Structural Model Exploration	0 Very Poor	Only one model was considered in paper.
Assumptions	1 Poor	Assumptions were listed. But consequences of assumptions were not explored (to our knowledge)
Parameter sensitivity	0 Very Poor	Sensitivity of predictions to parameter uncertainty not addressed (to our knowledge)

When employing receptor occupancy model, many assumptions are made

- The following phenomena are ignored.
  - Binding of target to its endogenous ligand
  - Feedback mechanisms for synthesis of target
  - Shedding and internalization of target
  - Tissue is assumed to be homogeneous
- Also specific assumptions are made about parameters
  - Tissue Biodistribution often is not known.

- The above assumptions may be ok, but how to check in a structured way?

We use model lumping to better understand system and improve Uncertainty Pedigree Table scores

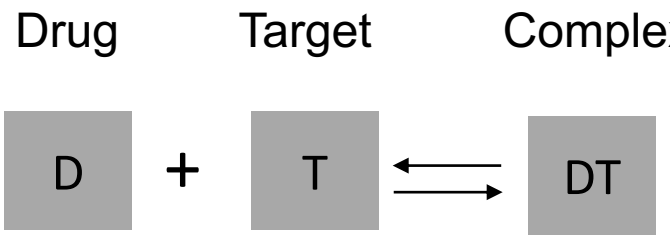
Automated Scale Reduction of Nonlinear QSP Models With an Illustrative Application to a Bone Biology System

Chihiro Hasegawa[1,2,*] and Stephen B. Duffull[1]

A combined model reduction algorithm for controlled biochemical systems

Thomas J. Snowden[1,2], Piet H. van der Graaf[3,2] and Marcus J. Tindall[1,4*]

In vitro, the target engagement depends only on drug conc and binding affinity



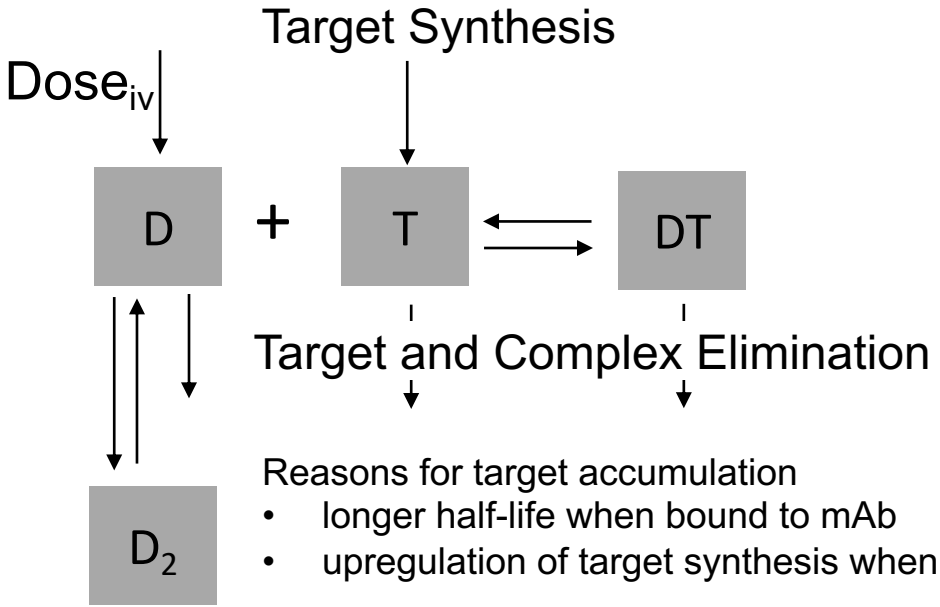
$$\text{Target Engagement} \approx \frac{C_{ss}}{C_{ss} + IC_{50}}$$

$$IC_{50} = K_{ss}$$

In vivo, accounting for turnover and accumulation of target may reduce engagement[1]

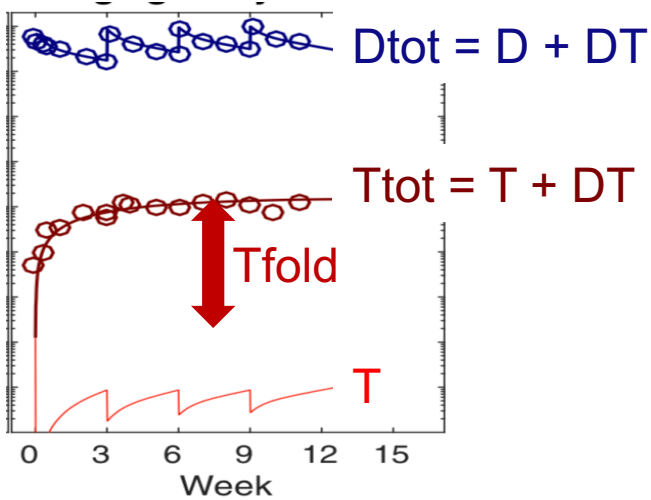
$$\text{Target Engagement} \approx \frac{C_{ss}}{C_{ss} + IC_{50}}$$

$$IC_{50} = K_{ss} \cdot T_{\text{fold}}$$

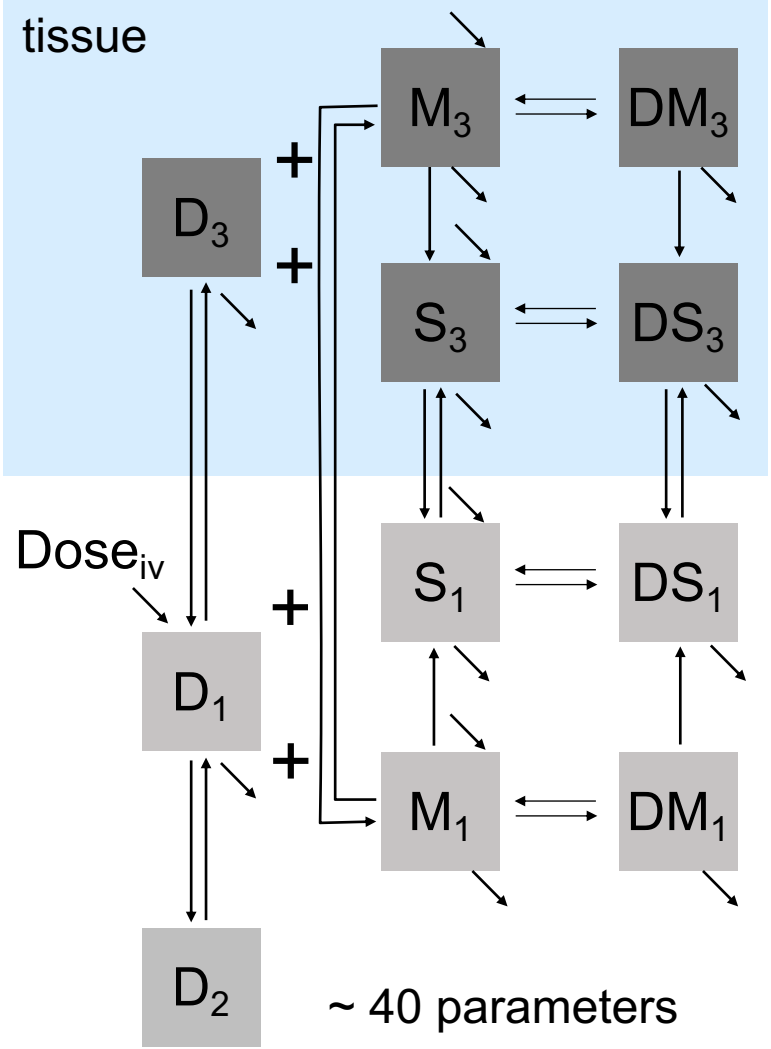


Reasons for target accumulation

- longer half-life when bound to mAb
- upregulation of target synthesis when inhibited



1. Stein AM, Ramakrishna R. AFIR: A dimensionless potency metric for characterizing the activity of monoclonal antibodies. Clin. Pharmacol. Ther: Pharmacometrics and Systems Pharmacol, 6, 258-266, 2017.



When shedding and tissue distribution are added, formula is similar[1]

$$\text{Target Engagement} \approx \frac{C_{ss}}{C_{ss} + IC_{50}}$$

$$IC_{50} = \frac{K_{ss} \cdot T_{\text{fold}}}{B}$$

B = biodistribution to tissue of interest

4 parameters or measured values

1. Ahmed S et al.,Guiding dose selection of monoclonal antibodies using a new parameter (AFTIR) for characterizing ligand binding systems. J Pharmacokin. And Pharmacodyn., 46.3, 297-304 2019.

List of assumptions to apply formulas:

- All systems
  - Target Engagement > 70%
  - Drug concentration >> Target concentration
  - Drug PK is linear
- When including tissue distribution
  - Tissue can be treated as homogeneous

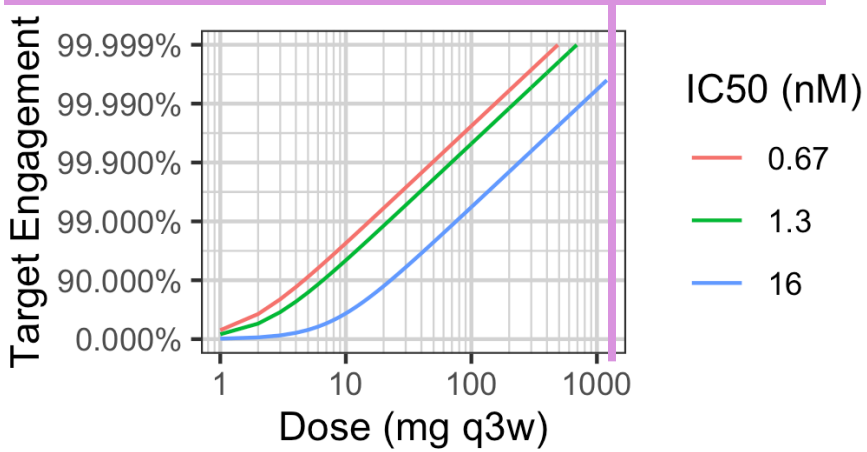
- This approach accounts for:
  - Shedding of target
  - Accumulation of target
  - Changes in target or ligand expression during therapy.

Parameter sensitivity for atezo shows 1200 mg q3w gives good coverage

	Best Case	Base Case	Worst Case
Css 5% ug/ml	50	50	50
Kss (nM)	0.4	0.4	0.8
B	0.3	0.3	0.1
Tfold	0.5	1	2
IC50(nM) Kss·Tfold/B	0.67	1.3	16

$$\text{Target Engagement} \approx \frac{C_{ss}}{C_{ss} + IC_{50}}$$

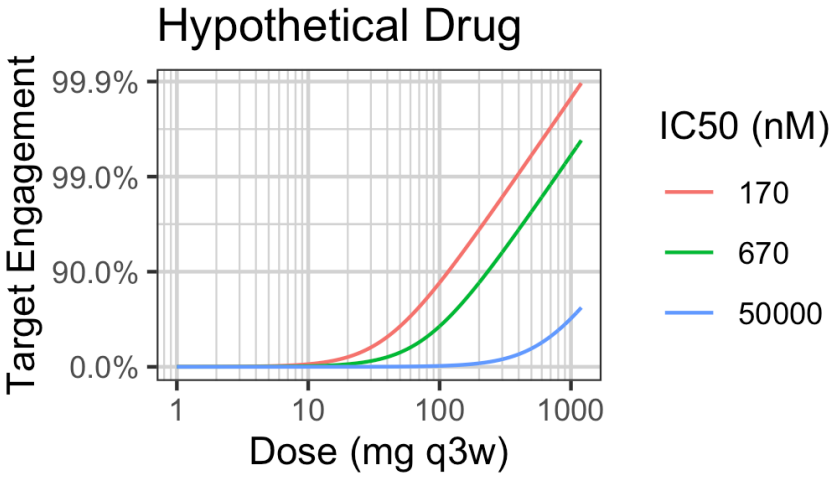
Atezo phase 2 dose (1200 mg) predicted to give high engagement



Parameter sensitivity for hypothetical drug shows assumptions are critical

	Best Case	Base Case	Worst Case
Css 5% ug/ml	20	20	20
Kss (nM)	1	1	5
B	0.3	0.3	0.1
Tfold	50	200	1000
IC50(nM) Kss·Tfold/B	170	670	50,000

$$\text{Target Engagement} \approx \frac{C_{ss}}{C_{ss} + IC_{50}}$$



Pedigree Assessment of original model

Criteria	Score	Description
Data	2 Fair	PK data is robust, Limited data on drug biodistribution to tumors
Biological Understanding	3 Good	General consensus on receptor binding kinetics, but receptor turnover and shedding was not addressed
Structural Model Exploration	0 Very Poor	Only one model was considered in paper.
Assumptions	1 Poor	Assumptions were listed. But consequences of assumptions were not explored.
Parameter sensitivity	0 Very Poor	Sensitivity of predictions to parameter uncertainty not addressed

Pedigree Assessment with analysis

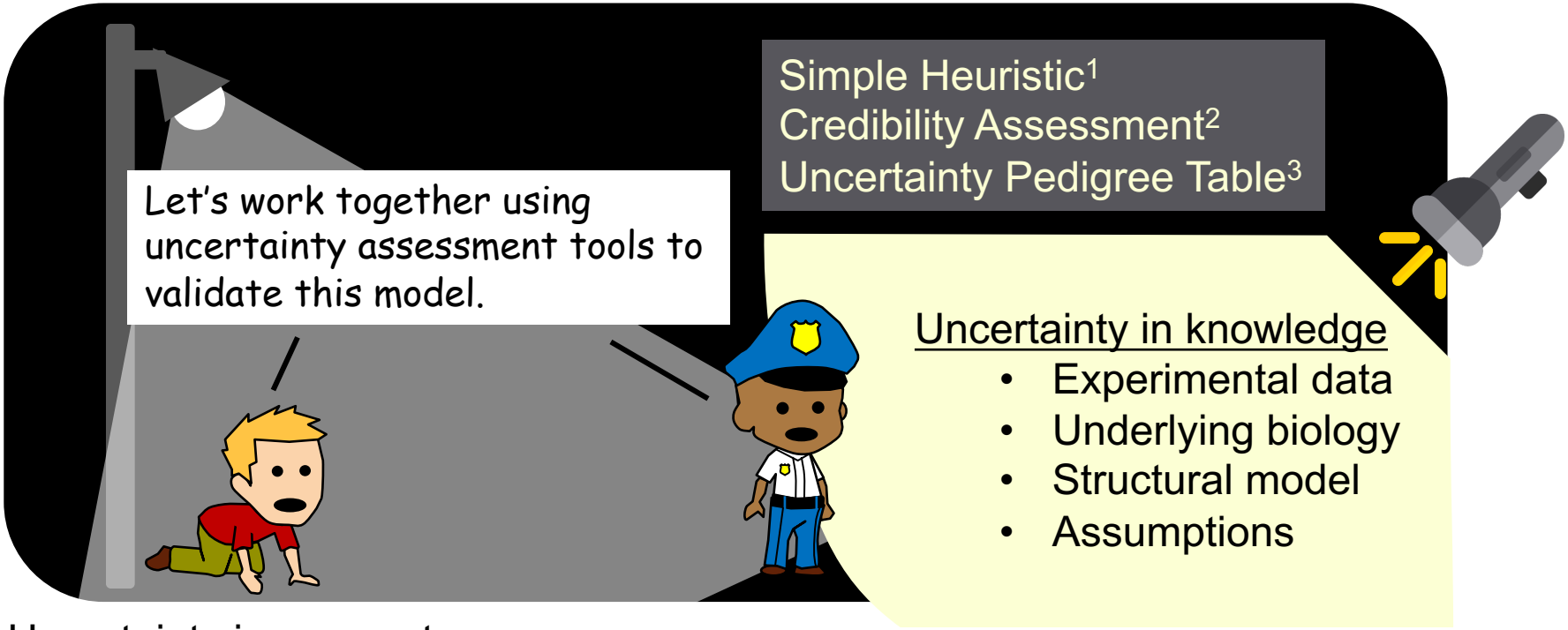
Criteria	Score	Description
Data	2 Fair	PK data is robust. Limited data on drug biodistribution and target accumulation inside tissue
Biological Understanding	3 Good	General consensus on receptor binding kinetics, but receptor turnover and shedding was not addressed
Structural Model Exploration	2 Fair	A set of models available for checking sensitivity
Assumptions	3 Good	Assumptions can be clearly listed and consequences explored through sensitivity analysis
Parameter sensitivity	3 Good	Sensitivity analysis performed

Application of receptor occupancy analysis

- Can guide us when to stop escalating a dose when drug is safe
  - Stop escalation when almost all patients are predicted to have very high target engagement
- Can help suggest dose(s) that should be tested in Phase 2
  - Merck used this approach for Pembrolizumab

- It's always preferable to use more downstream PD endpoints, but sometimes (e.g. atezolizumab) these biomarkers (e.g. tumor size) may not be available until the pivotal study.

Streetlight effect in model assessment

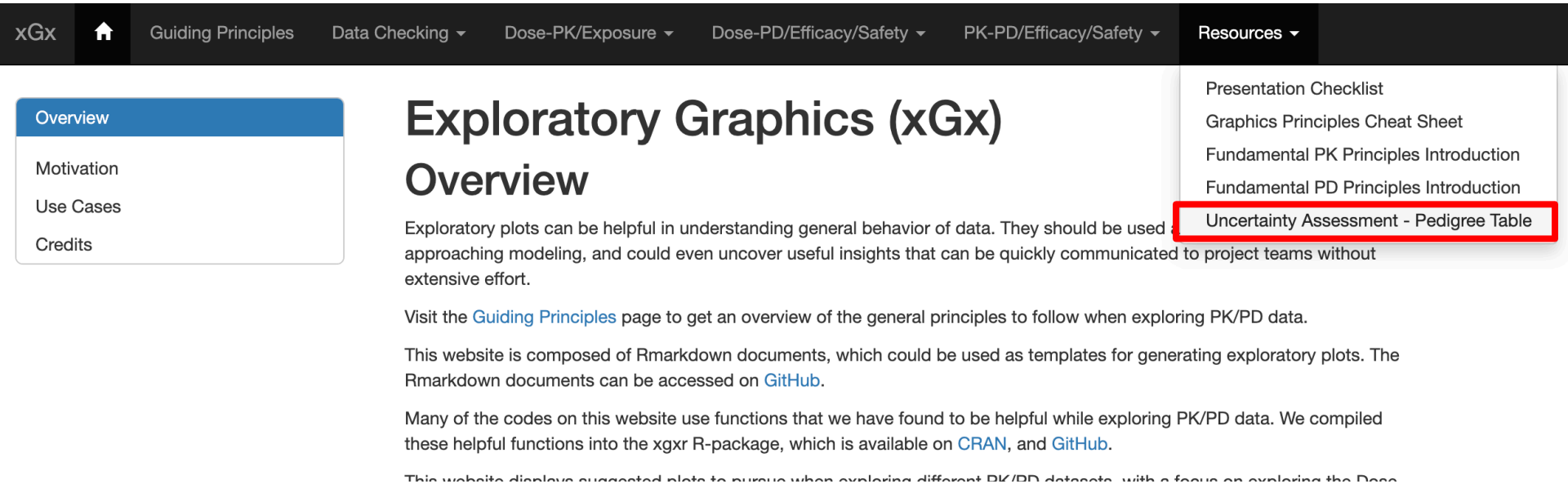


Uncertainty in parameters

- Sensitivity Analysis
- Virtual Patient Simulation

1. Stein and Looby, et al., CPT:PSP, 2018, doi:10.1002/psp4.12311
2. Kuemmel, Colleen, et al. CPT:PSP, 2019 doi:10.1002/psp4.12479.
3. https://opensource.nibr.com/xgx/Resources/Uncertainty_Assessment_Pedigree_Table.pdf

Pedigree table available at: opensource.nibr.com/xgx



This is a work in progress. We are currently trying this approach out in case studies

Executive Summary

- More QSP models are being submitted to the FDA.
- How to validate these models?.
- Three tools have been presented for evaluating QSP models
  - Simple Heuristic (when available)
  - Context of Use Table
  - Uncertainty Pedigree Table

Acknowledegements

- Andrea Saltelli[1-2]
- Pharmacometics and PKS-Modeling Groups
  - Tycho Heimbach
  - Jeff Kearns
  - Jaeyeon Kim
  - Mick Looby
  - Alison Margolskee
  - Birgit Schoeberl

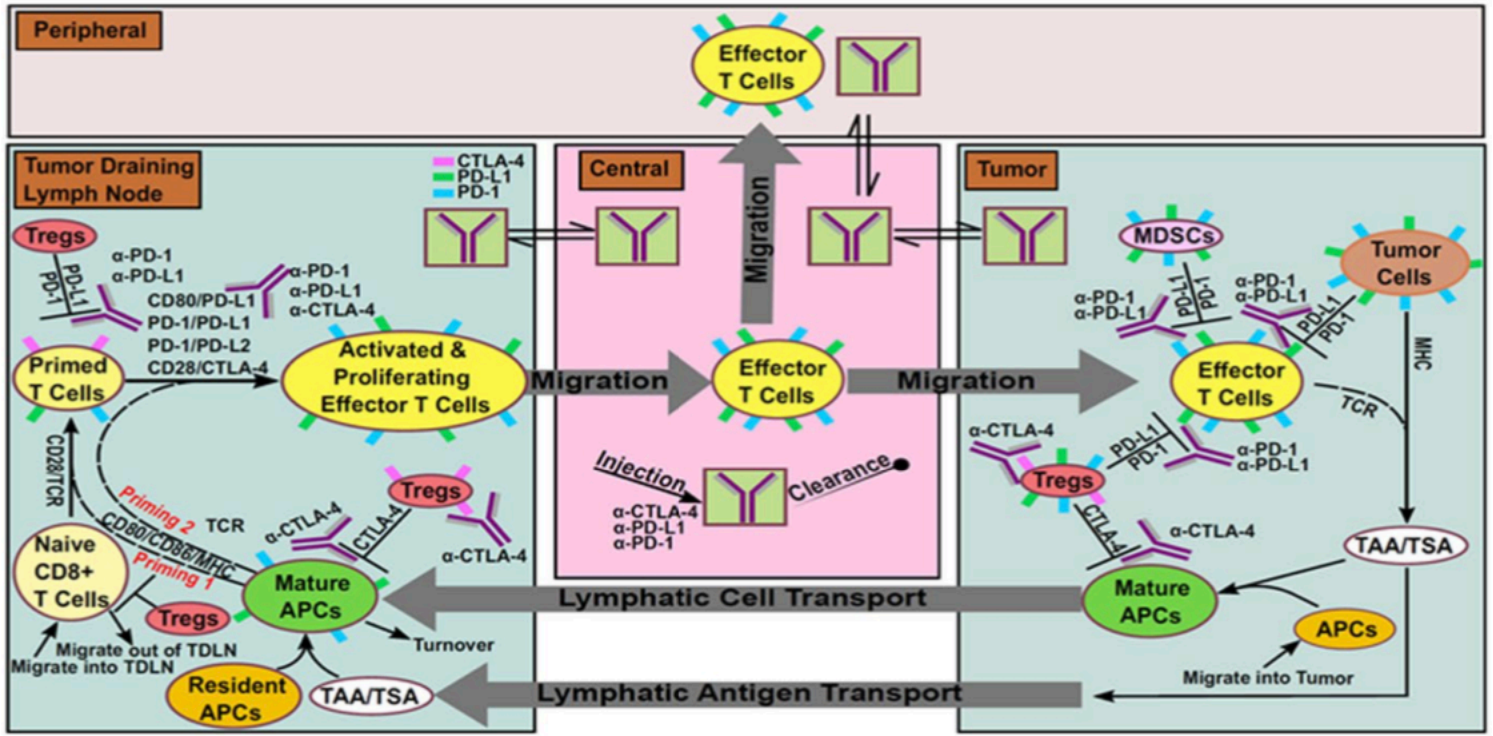
1. Saltelli, Andrea. "A short comment on statistical versus mathematical modelling." Nature communications 10.1 (2019): 1-3.
2. Saltelli, Andrea, et al. Sensitivity analysis in practice: a guide to assessing scientific models. Vol. 1. New York: Wiley, 2004.

Backups

Repeated Predictions: Compare QSP to Simple Heuristics

Stein AM, Looby M. Benchmarking QSP models against simple models: A path to improved comprehension and predictive performance. *Clin. Pharmacol. Ther: Pharmacometrics and Systems Pharmacol,* 7.8, 487, 2018.

Model for immuno-oncology



Many questions remain about how to predict immune response

- How many cell types need to be modeled?
- Do we even know all cell types that need to be included in the model?
- Is it reasonable to treat cell types as discrete populations, or do they need to be treated as a continuum?
- How many cytokines and which effects on which cell types need to be modeled?
- What parts of the tumor micro-environment need to be considered?
- Does the tumor need to be treated as a spatially heterogeneous structure?

- Benchmarking any predictive QSP model against a simple heuristic would help in validating the model and evaluating the bias-variance trade-off.

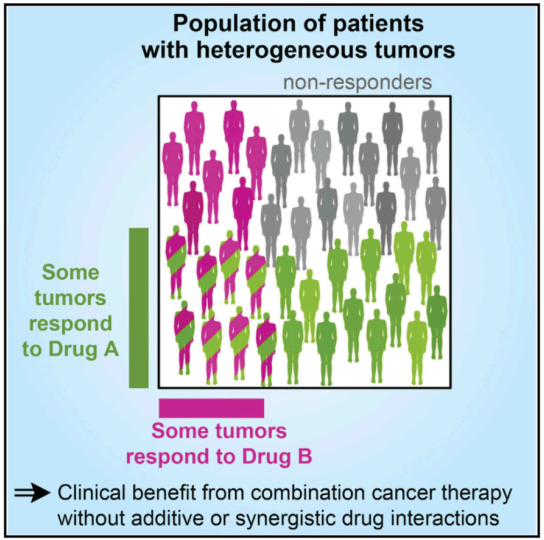
Simple heuristic available for predicting combo response from monotherapy

Theory

Cell

Combination Cancer Therapy Can Confer Benefit via Patient-to-Patient Variability without Drug Additivity or Synergy

Graphical Abstract

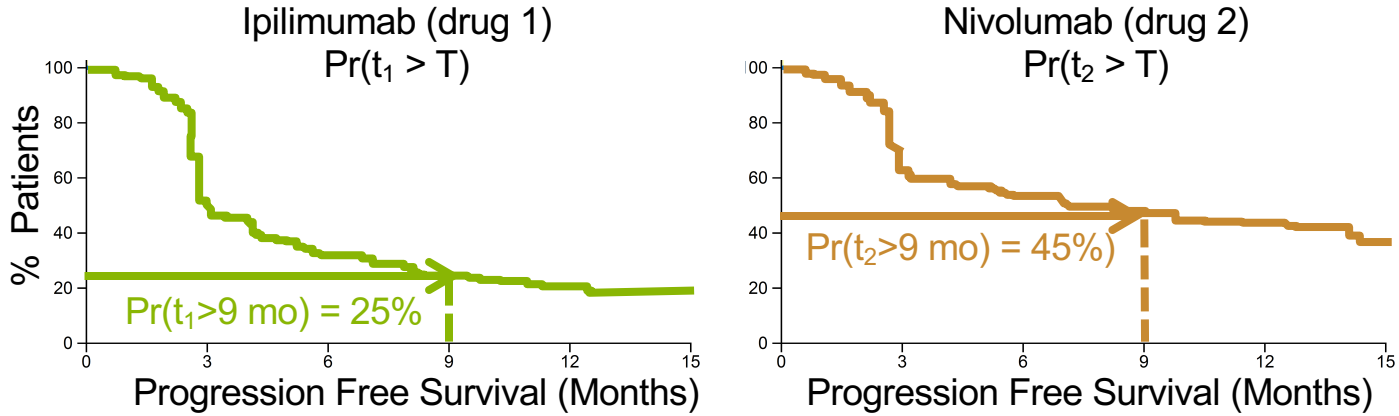


Authors
Adam C. Palmer, Peter K. Sorger

Correspondence
peter_sorger@hms.harvard.edu

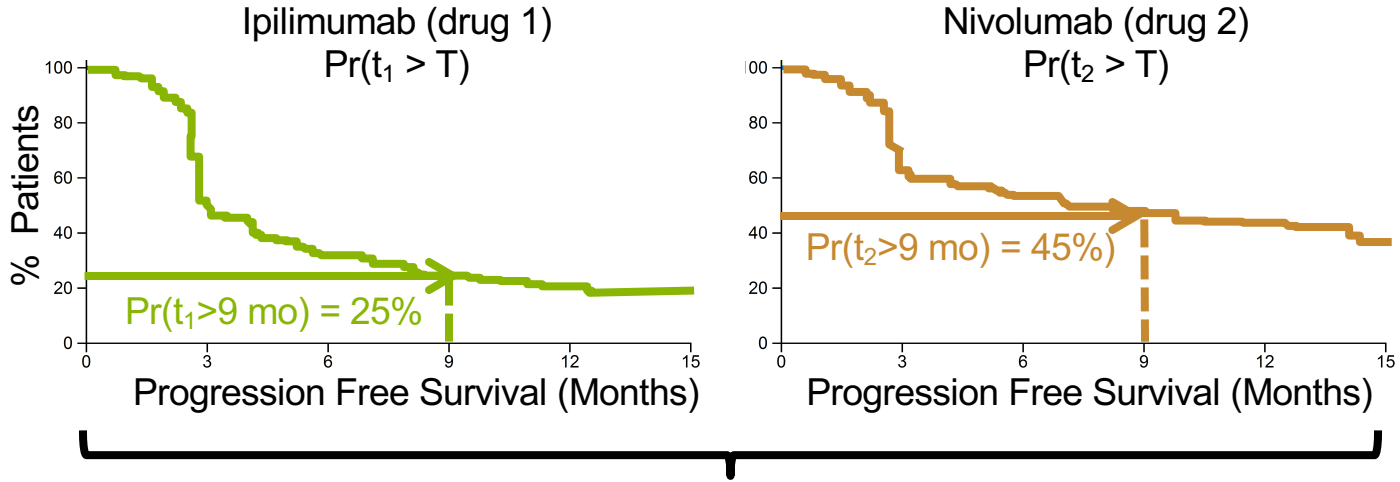
In Brief
Patient-to-patient variability in response to single drugs is sufficient to explain the efficacy of a large number of combination cancer therapies without pharmacologically additive or synergistic effect in individual patients.

Assume “independence of action” of therapies



Larkin *et al* (2015) *New England Journal of Medicine* **373**:23

Assume “independence of action” of therapies predicts combo PFS



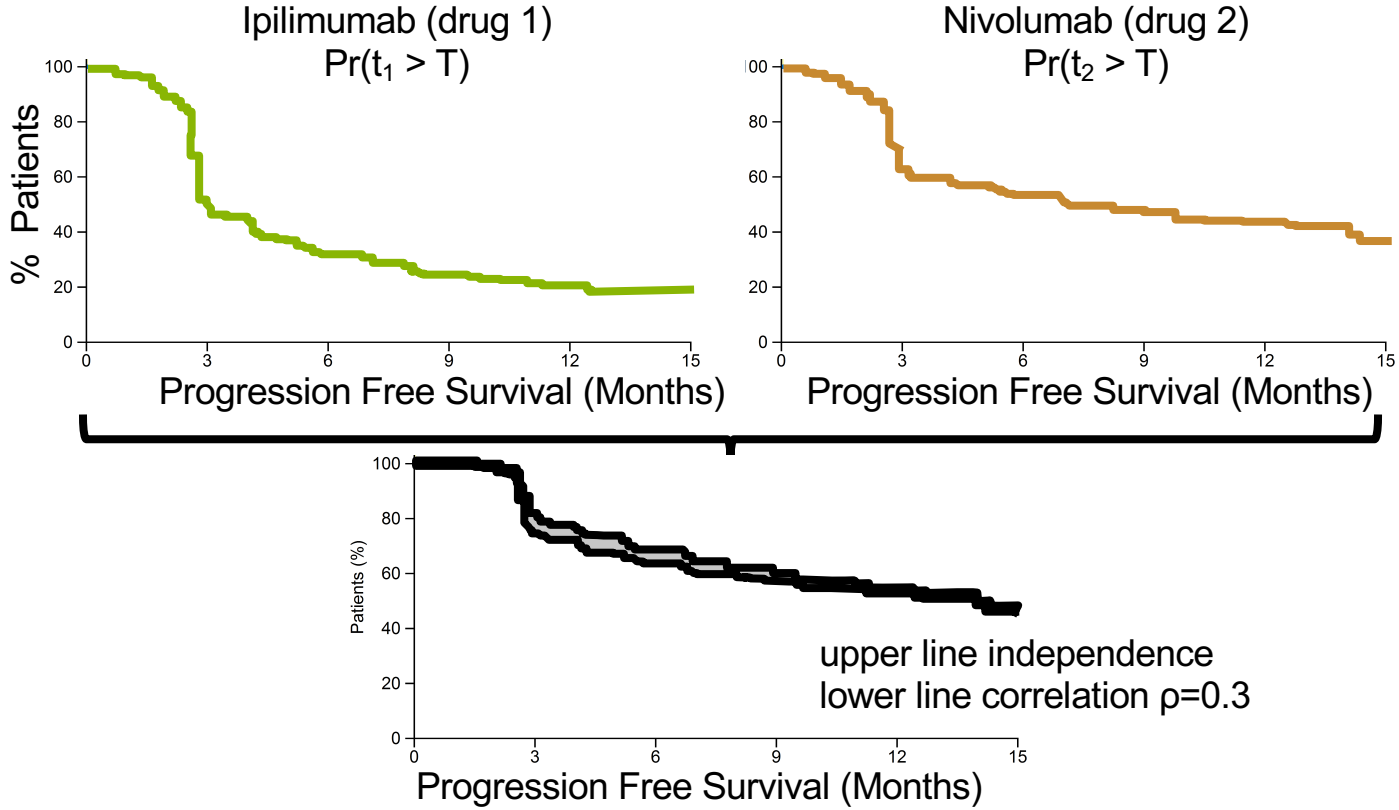
For combo: $Pr(t_{1+2}>T) = Pr(t_1>T \text{ or } t_2>T)$

Assuming independence: $Pr(t_{1+2}>T) = Pr(t_1>T) + Pr(t_2>T) - Pr(t_1>T)\cdot Pr(t_2>T)$

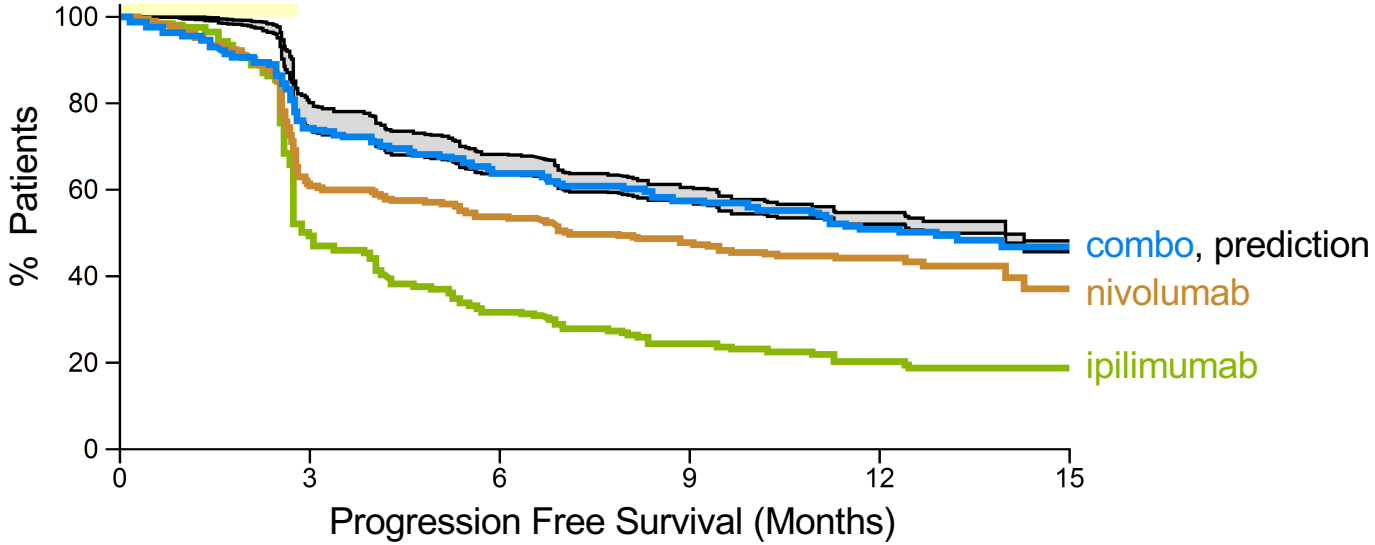
$Pr(t_{1+2}> 9 \text{ mo}) = 25\% + 45\% - 25\% \cdot 45\%$

$= 59\%$

Assume “independence of action” of therapies predicts combo PFS



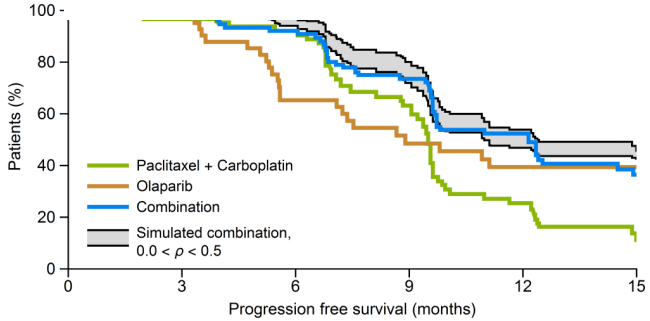
Independent activity predicts ipilimumab + nivolumab combo



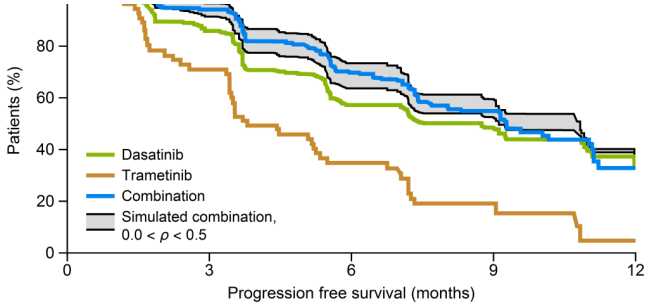
Larkin *et al* (2015) *New England Journal of Medicine* **373**:23

Two thirds of trials could be predicted by “independent” activity

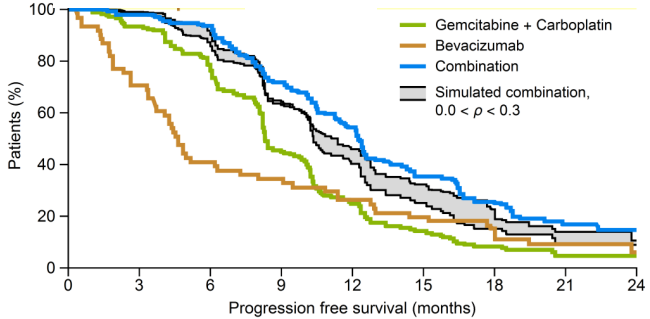
Recurrent ovarian cancer



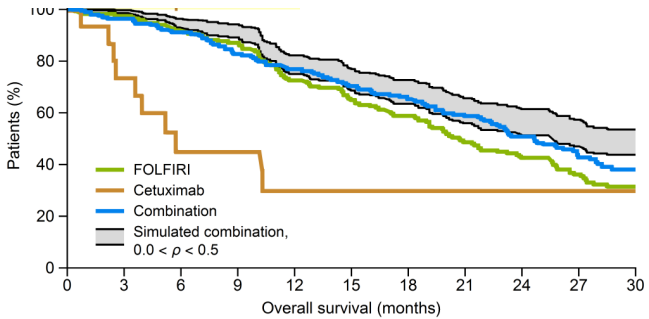
BRAF-mutant melanoma



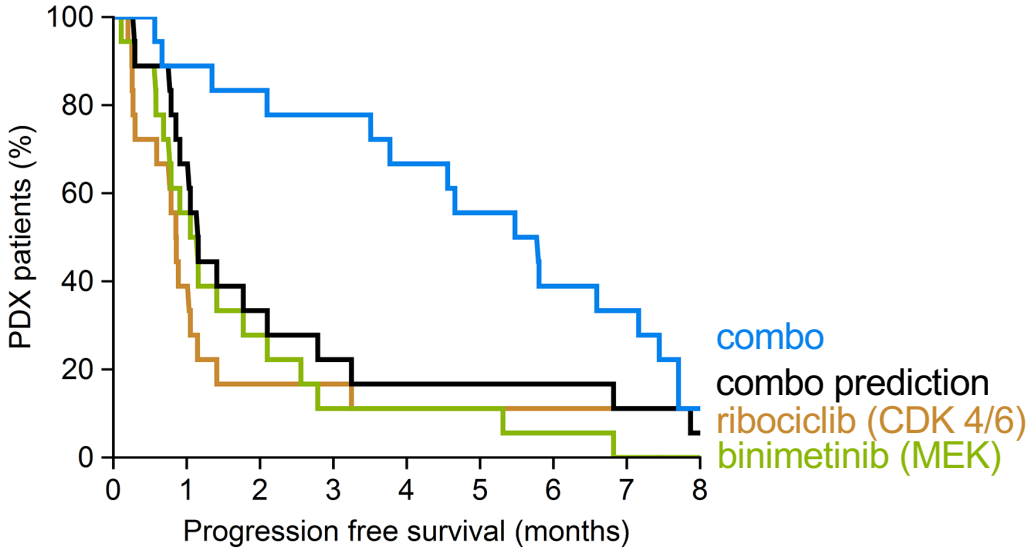
Recurrent ovarian cancer



KRAS-wildtype colorectal cancer



Exception: synergy example in patient derived xenograft for melanoma



Analysis of data from Gao *et al* (2015) *Nature Medicine* **21**:1318

Pedigree Table

Uncertainty pedigree table scores are qualitative and may vary by reviewer

- Different reviewers have different scores
- This difference is ok. The most important thing is to have the discussion.
- Difference in opinion could be useful to report and explore further

Criteria	Modeler	Reviewer1	Reviewer2
Data	4	2	4
Biological Understanding	3	1	4
Structural Model Exploration	3	2	2
Assumptions	3	3	3
Parameter sensitivity	4	4	4

Van Der Sluijs, JP,
Risk Analysis: An International Journal 25.2 (2005): 481-492.

1. Experimental data quality

Criteria	Score	Assessment
Very Good	4	All data from validated assays and reproducible experiments. Data fully describes system
Good	3	Enough trusted data for good description of system
Fair	2	Limited data with old experiments using unvalidated assays
Poor	1	Very limited data, educated guesses inform the model
Very Poor	0	No data

2. Consensus on underlying biology and its mathematical description

Criteria	Score	Assessment
Very Good	4	Well-established theory
Good	3	Accepted theory
Fair	2	Limited consensus
Poor	1	Embryonic field
Very Poor	0	Crude speculation

3. Structural model exploration

Criteria	Score	Assessment
Very Good	4	Models representing all known sources of uncertainty
Good	3	Many structural models, including random model perturbations
Fair	2	A few structural models representing key sources of uncertainty
Poor	1	A few structural models were considered
Very Poor	0	Only a single structural model considered

4. Clear statement and understanding of the assumptions

Criteria	Score	Assessment
Very Good	4	Few assumptions, clearly listed. Impact of assumptions is clear
Good	3	Moderate assumptions, clearly listed. Impact is clear
Fair	2	Many assumptions, clearly listed. Impact was explored but is difficult to understand
Poor	1	Assumptions are listed, but impact not explored or explained
Very Poor	0	Assumptions are not clearly listed

5. Accurate interpolation

Criteria	Score	Assessment
Very Good	4	Model fits validated using standard diagnostics (VPCs, residuals)
Fair	2	Model fits consistent with data
Very Poor	0	Not designed for interpolation

6. Accurate extrapolation

Criteria	Score	Assessment
Very Good	4	Many extrapolated predictions were confirmed
Fair	2	Moderate confidence in ability to extrapolate
Very Poor	0	Not designed for extrapolation

7. Intersubject variability

Criteria	Score	Assessment
Very Good	4	The key sources of variability are well understood and included
Good	3	Reasonable hypotheses exist for main sources of variability
Poor	1	Many guesses made for where and how to include variability
Very Poor	0	Variability is not included in the model.

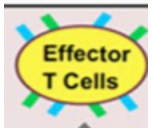
8. Parameter sensitivity and identifiability

Criteria	Score	Assessment
Very Good	4	Global analysis over well-informed range or established theory exists for understanding system
Good	3	Global analysis with minor doubts of realism of uncertainty distribution and choice of fixed parameters
Fair	2	Global analysis, with significant doubts of realism of assumed uncertainty distribution and choice of fixed parameters
Poor	1	Local analysis only
Very Poor	0	No

Quality control

Criteria	Score	Assessment
Very Good	4	Validated, reproducible computing environment with a careful audit of the most likely sources of errors in programming by an independent modeler
Fair	2	Validated reproducible, computing environment, modeler performed self-audit
Very Poor	0	Unvalidated computing environment, no audit

Many uncertainties in how to model effector T cells



- Rules governing when T cells stop proliferating and become exhaustion
- Heterogeneity in T cell receptor clones
- Heterogeneity in degree of expansion of identical clones
- Rules governing cycling between effector and memory cells
- How many different cell-surface markers to include?

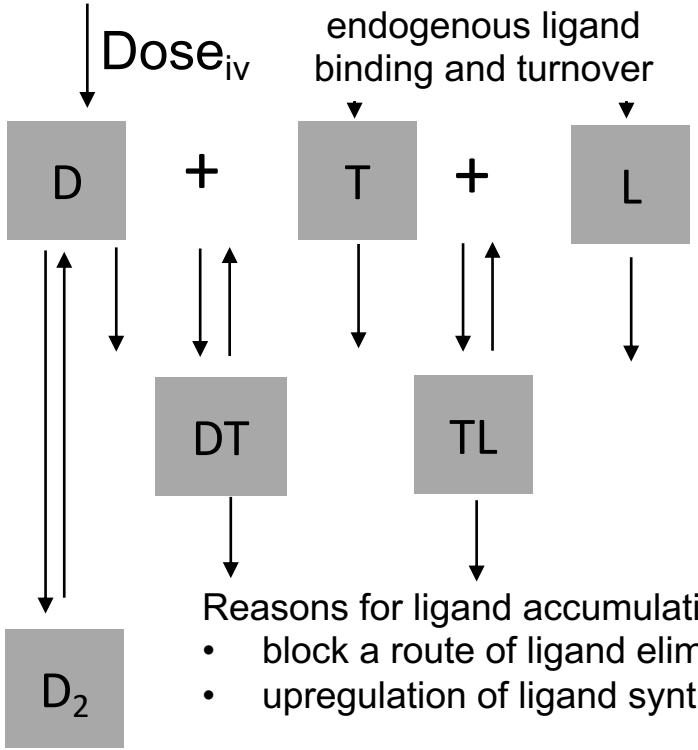
In many fields, simple models outperform complex models[1-2]

- Earthquakes
- Sports
- Criminology
- Stock Market
- Sales

1. Silver, Nate. *The signal and the noise: why so many predictions fail--but some don't*. Penguin, 2012.
2. Haldane, Andreiv G., and Vasileios Madouros. "The dog and the frisbee." Revista de Economía Institucional 14.27 (2012): 13-56.

Lumped Parameters

When endogenous ligand is present, it can accumulate, too[1]

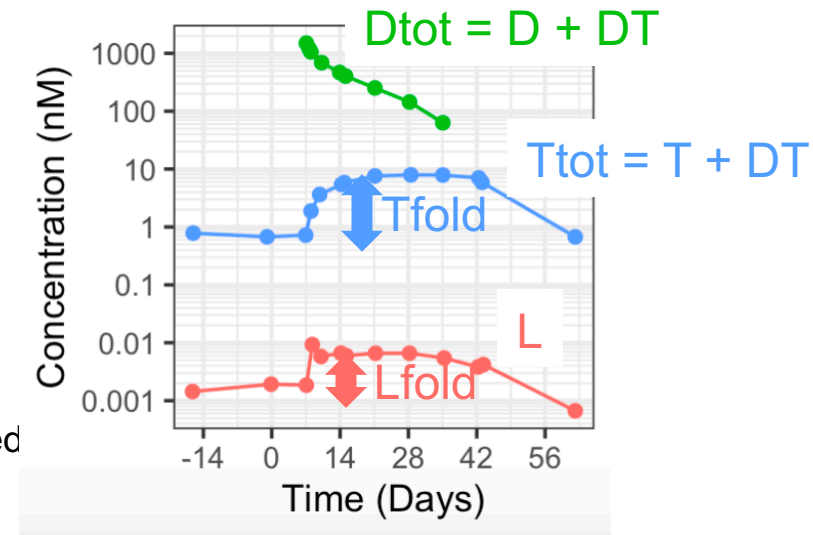


Reasons for ligand accumulation

- block a route of ligand elimination
- upregulation of ligand synthesis when inhibited

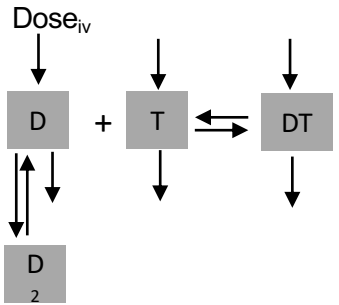
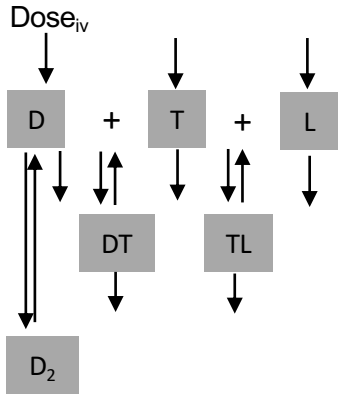
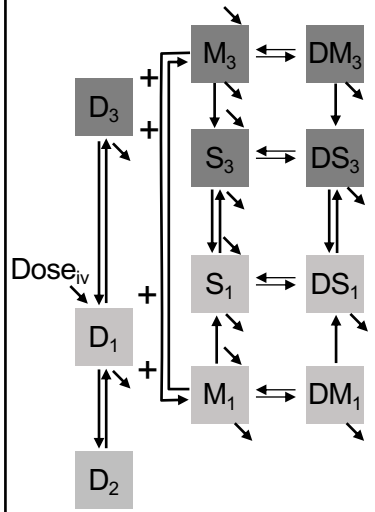
$$\text{Target Engagement} \approx \frac{C_{ss}}{C_{ss} + IC_{50}}$$

$$IC_{50} = K_{ss} \cdot T_{\text{fold}} \cdot L_{\text{fold}}$$



1. Stein AM et al. Biorxiv (to be submitted)

$$\text{Target Inhibition} = \frac{C_{ss}}{C_{ss} + IC_{50}}$$

Name	Binding	TMDD	TMDD+L	TMDD+tissue
Tissue	In vitro	Circulation	Circulation	Tissue
Ligand	no	no	yes	no
IC_{50}	K_{ss}	$K_{ss} \cdot T_{fold}$	$K_{ss} \cdot T_{fold} \cdot L_{fold}$	$\frac{K_{ss} \cdot T_{fold}}{B}$
Model				

D = Drug
T = Target
L = Ligand
M = Membrane-bound target
S = Soluble target (shed)

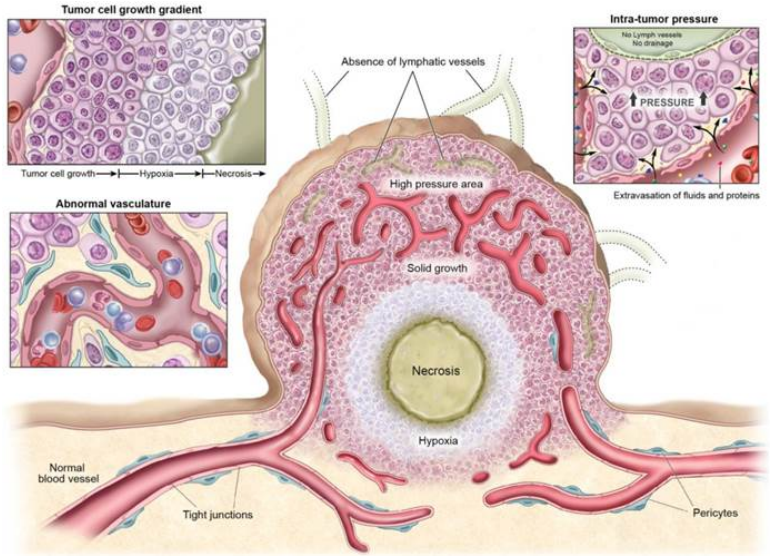
1 = central compartment
2 = peripheral compartment
3 = tissue compartment

T_{fold} = fold-increase in total target after binding drug
L_{fold} = fold-increase in ligand after drug binds target
B = tissue biodistribution coefficient

Intratumor Heterogeneity

Tumors are not well-mixed systems

- Uneven vascularity
- Poor lymphatic drainage
- High interstitial pressure



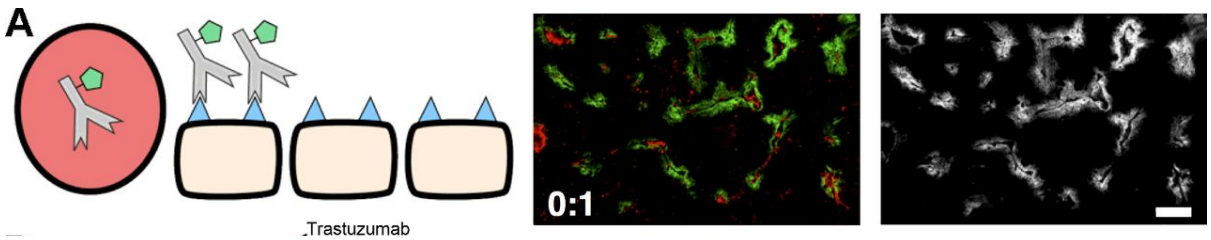
http://www.thno.org/v04p0081.htm

Diffusion (concentration gradient) dominates convection (fluid flow)

Heterogeneous distribution of mAbs

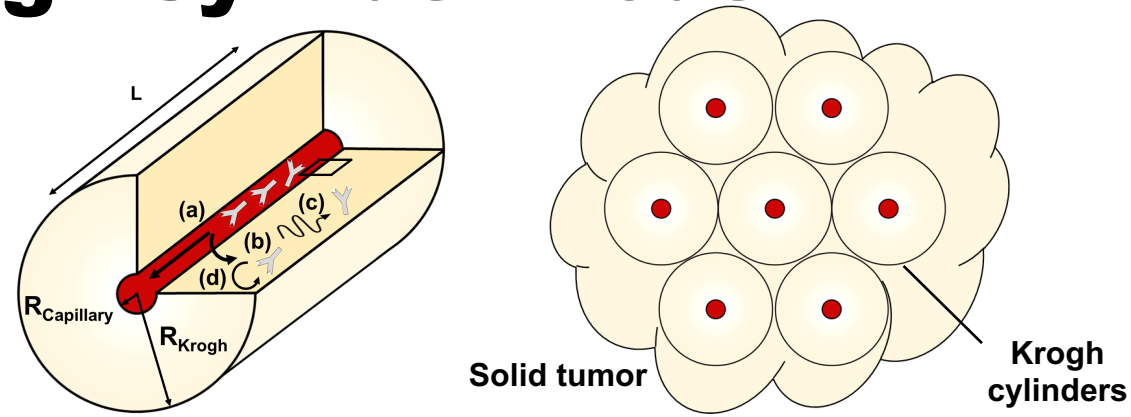
- Competition between binding and diffusion creates the “binding-site barrier”
- Administration of very high doses can overcome this barrier

What constitutes a “high dose”?



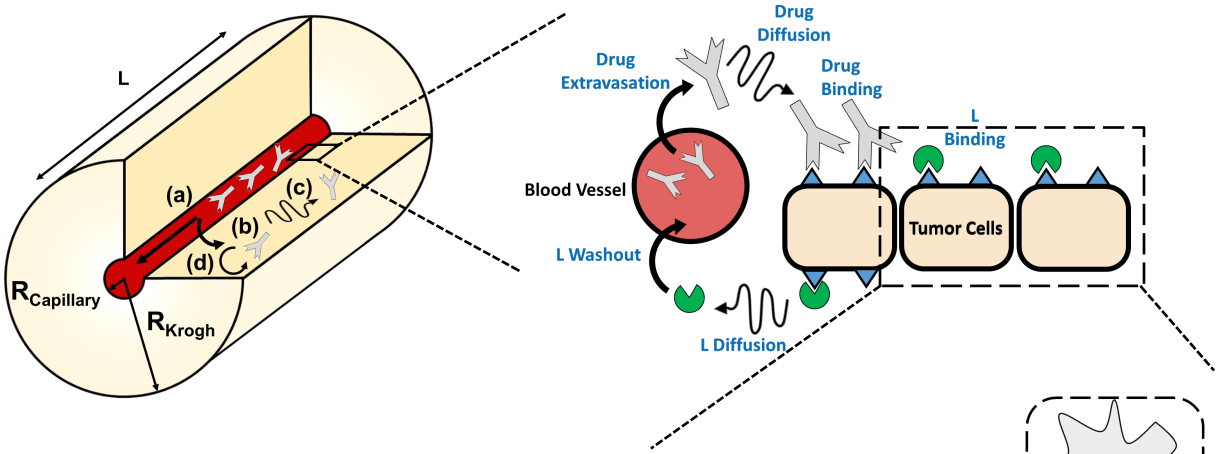
Cilliers et al. *Cancer Res*; (2018) 78(3); 758–68

Krogh Cylinder Model



- Krogh cylinder is unit element of a tumor
- Four main processes a drug undergoes
  - (a) Blood flow
  - (b) Extravasation (permeability-dependent uptake)
  - (c) Diffusion (mass transport in extracellular space)
  - (d) Local metabolism (cellular internalization and degradation)

Mechanistic Tumor Model



- Systemic pharmacokinetics
- Intra-tumoral distribution
- Cellular metabolism