Collaborations with a Biostatistician: From clinical trial designs to translational research with high-dimensional data

Kim-Anh Do, Ph.D.
Professor and Chair
Department of Biostatistics
kimdo@mdanderson.org

Faculty

Department of Biostatistics

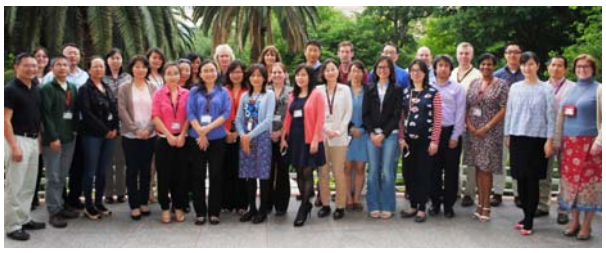
Faculty

22 faculty members (16 tenured/tenure-track)

- Collaborate with and supervise the work of
  - analysts & research statisticians
  - postdoctoral fellows and graduate research assistants
- Collaborate with scientists and clinicians
  - at MD Anderson
  - at other institutions – national and international collaborations

Statistical Analysts

- Collaborate with researchers on study design
- Prepare statistical considerations for grant applications
- Perform data analysis services



- Participate in review of research protocols for statistical strength
- Staff drop-in statistics clinics three times per week
- Work with Biostatistics faculty on research projects/collaborations

MD Anderson

Quantitative Research Computing Group

Department of Biostatistics

Team of programmers and systems analysts conducting:

- Information computing services
- Data management initiative
- BCB software engineering
- Bayesian numerical analysis and solutions

BCB Software Engineering

Clinical Trial Design and Implementation Services

Information and Computing Services

Data Management Initiative Group

Vision

Develop and lead the Department of Biostatistics as a nationally and internationally recognized and well-funded program

Translate genomic knowledge and technology to bridge the gaps between

- biostatistics
- bioinformatics
- computing
- applications to biology, clinical medicine, public health
- benefits to our patients and society at large

Develop a well-rounded department that is world-famous for

- Research — novel methodology & collaborative endeavors
- Training postdoctoral fellows, graduate students, and analysts

Objectives

- Develop innovative statistical methods with application to cancer research
- Provide biostatistical collaboration, consultation, and quantitative research resources to clinical, laboratory, and prevention scientists for the planning, conduct, analysis, quality assurance, and interpretation of research studies
- Train the next generation of biostatisticians

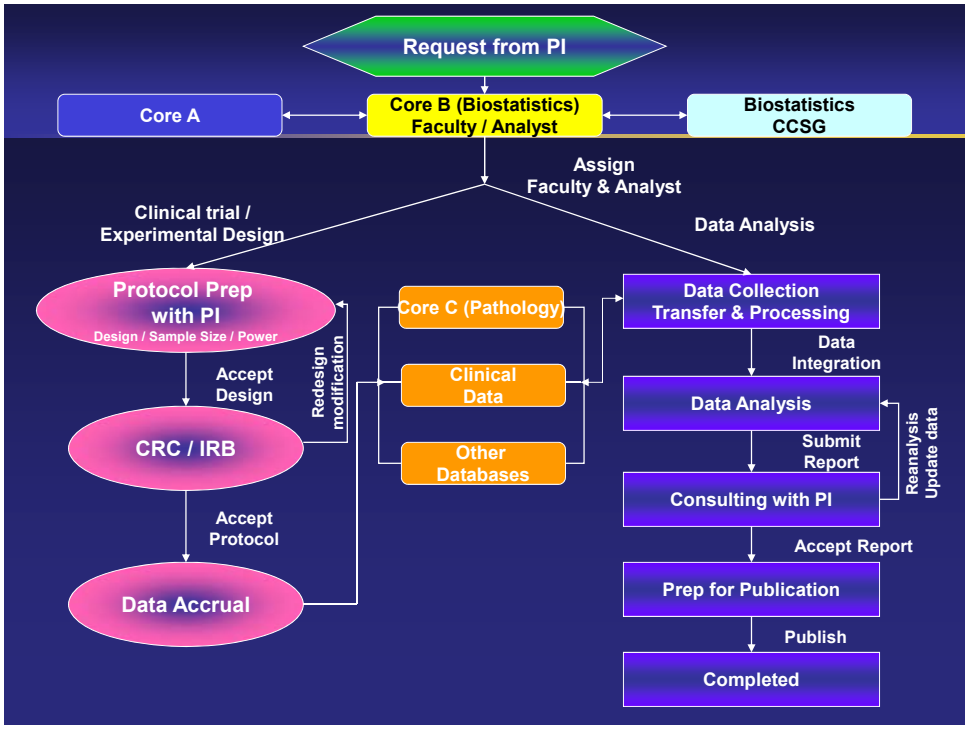
Statistical applications in clinical, basic science, and behavioral/social science research

- Integrated statistical modeling of high-dimensional biomarker and complex functional, and imaging data
- Computer-intensive statistical methods
- Bayesian modeling, computation, and inference
- Clinical trial design and analysis (e.g., adaptive randomization)
- Cancer screening and early detection research
- Sequential statistical designs
- Statistical modeling for behavior and social science applications

Impactful Research

Five impactful areas of research:

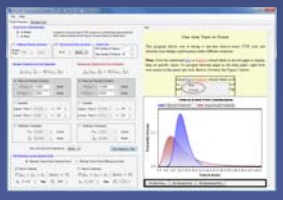
- Integrative analysis of multiplatform genomic and other high-dimensional data
- Early cancer detection, statistical & data coordinating center
- Imaging statistics
- Innovative clinical trial design and analysis
- Behavioral and social statistics



Innovative Clinical Trial Design and Software Services

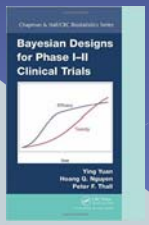
CTC:

This group provides a Clinical Trial Conduct (CTC) software system where clinicians use a web-based user interface applying Bayesian statistical methods to the conduct of clinical trials. Over 290 trials have been conducted by 700 users at multiple institutions using a variety of innovative statistical methods.



Recent Windows software releases:

- One Arm Time-To-Event monitoring program
- BMA-CRM dose-finding program updates
- Eff-Tox dose-finding program updates
- BOIN (Bayesian Optimal Interval design) program
- Numerical Software Library (C++) updates



Recent book publication:

- "Bayesian Designs for Phase I-II Clinical Trials", June 2016, with supporting C++ software applications

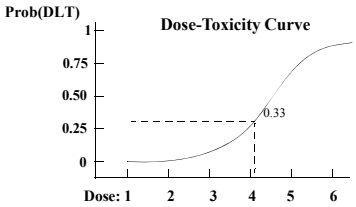
Innovative Clinical Trial Design and Software Services

This group is also known as the Numerical-Statistical Software Group. It provides innovative clinical trial software design services, including software and statistical tools for all stages of clinical trial design, planning, conduct, and analysis.

Functions:

- Collaborate with other researchers to develop software for innovative trial designs, often published in prominent peer-reviewed statistical journals.
- Maintain and support core statistical software applications produced via modern software engineering processes. Produce, maintain, and monitor an infrastructure supporting software development with 24/7 system availability.
- Develop and maintain a website (CTC) used world-wide by clients for calculating sophisticated statistics for conducting clinical trials. Provide user management, security, validated numerical calculations.
- Provide a Software Download Kiosk (website) from which any statistician can download over 85 software packages authored by this group. Software has been downloaded by 17,500 scientists worldwide.

Design and Statistical Considerations for Phase I Cancer Clinical Trials



1. Premises:
   There is an unknown dose-toxicity curve.
   Higher toxicity is associated with higher efficacy.
   Given a certain target toxicity level (TTL) one is willing to accept, the goal is to find the maximum tolerated dose (MTD) which yields the TTL.

2. Key Elements
   (a) Define the dose limiting toxicity (DLT)
   (b) Choose a starting dose
   (c) Define the dose escalation scheme in terms of dose spacing, dose assignment, and cohort size
   (d) Determine the MTD by an algorithm or a model-based method

3. Major drawback of the conventional Phase I design:
   starts too low, escalates too slow, and takes too long

Dose Assignment / Cohort Size of the Conventional 3+3 Design

1. Enter 3 patients at the starting (lowest) dose level

2. If 0/3 has DLT → Next 3 pts at the next higher dose level

   If 1/3 has DLT → 3 more pts at the same dose level

   1/3 + 0/3 has DLT → Next 3 pts at the next higher dose level
   1/3 + ≥ 1/3 have DLT → Exceeds the MTD

   If 2/3 or 3/3 have DLT → Exceeds the MTD

3. If the current dose has not exceeded the MTD, repeat step 2

4. If the current dose has already exceeded the MTD:
(a) current dose = starting dose: MTD not defined
(b) current dose > starting dose
(i) only 3 pts treated at the previous level, enter 3 more pts at that level 0/3 + 0-1/3 DLT → Declare the level as MTD
0/3 + 2-3/3 DLT → Exceed the MTD go back to step 4

(ii) 6 pts already treated at the previous level (must be 1/6 DLT):
→ Declare that level as MTD

5. If need to escalate beyond the last dose level: MTD not defined

One drawback: The MTD does not correspond to a fixed TTL. For example, the 3+3 Design may choose the MTD at a level of 23% DLT in one trial, and in another trial, it may choose the MTD at a level of 15% DLT.

Another drawback: The algorithm is not flexible enough to choose the MTD at any TTL. In practice, different TTLs may be accepted depending on the patient population, disease site, available trtmt

Examples

1.

Cohort	Dose Level 1	2	3	4	5
1	0/3				
2		0/3			
3			1/3		
4			0/3		
5				2/3	
MTD			***		

2.

Cohort	Dose Level 1	2	3	4	5
1	0/3				
2		0/3			
3			0/3		
4				2/3	
5			0/3		
MTD			***		

3.

Cohort	Dose Level 1	2	3	4	5
1	0/3				
2		0/3			
3			0/3		
4				2/3	
5			2/3		
6		1/3			
MTD		***			

Continual Reassessment Method (CRM)

(O'Quigley et al, Biometrics, 1990)

1. A Bayesian methodology to estimate the dose-toxicity curve and to assign patient at a level closest to the current estimate of the MTD

2. Choose a family of dose-toxicity curve to model
***p* = Prob(DLT at dose *d*) = *f* (*d* | *a*) with parameter *a*, e.g.:**

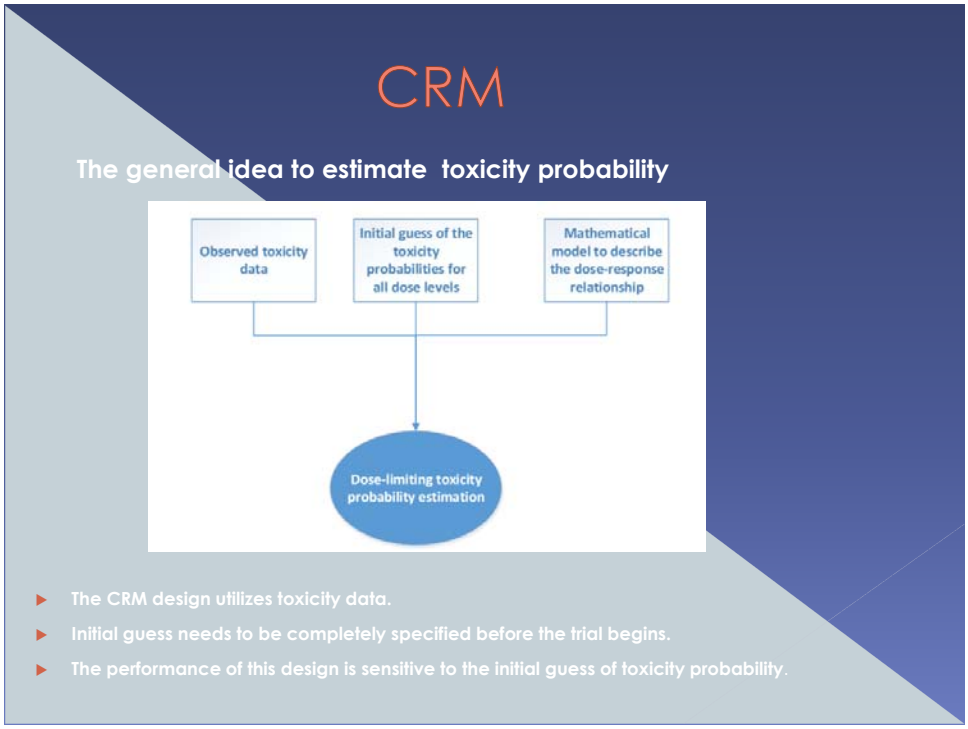
Hyperbolic tangent model: $p = \left[\frac{\tanh(d)+1}{2}\right]^a = \left[\frac{e^d}{e^d+e^{-d}}\right]^a$

One-parameter logistic model: $p = \frac{e^{3+a\cdot d}}{1+e^{3+a\cdot d}}$

Power model: $p = d^{\exp(a)}$

3. **Goal: find the MTD *d**** which yields the pre-specified TTL ***p****
i.e.: $d^* = f^{-1}(p^* \mid a)$

4. Scheme
(a) Assume a vague or non-informative prior distribution for *a*.
(b) Given the current information of *a*, treat 1 patient at the dose level closest to the current estimate of the MTD
(c) Observe the toxicity outcome of the patient
(d) Update the info of *a* by computing its posterior distribution
(e) Repeat steps (b)-(d) until the max number of patients is reached.



Properties of CRM

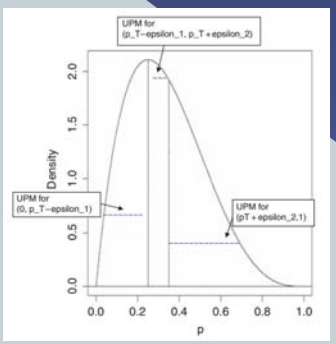
Advantages of CRM

1.It is a model-based method with a clearly defined objective.
2.It treats more patients at doses closer to the target MTD and, hence, reduces the number of patients treated at ineffective dose levels.
3.Because more patients are treated around the target level, it gives more precise estimate of the Prob(tox) at MTD.
4.Although CRM does not determine the lowest dose level and dose spacing, there is no restriction of treating patients only at pre-speci- fied doses. (e.g.: continuous infusion)

Disadvantages of CRM

1.It may be TOO AGGRESIVE. Patients may be treated at very high toxic doses in the early phase of the trial.
2. Success depends on the proper choice of the dose-toxicity curve and the prior distribution of its parameter(s).
3.Because of cohort size 1, it may take long time to complete a trial with high patient accruals.
4. Need special computer programs to implement the design.

mTPI (modified Toxicity Prob Interval Design)



p_T is target toxicity prob
Only requires a definition of an equivalence interval (EI)

$[p_T - \varepsilon_1,\ p_T + \varepsilon_2]$

Physician specifies lowest toxicity prob acceptable to treat future patients without dose escalation → $p_T - \varepsilon_1$

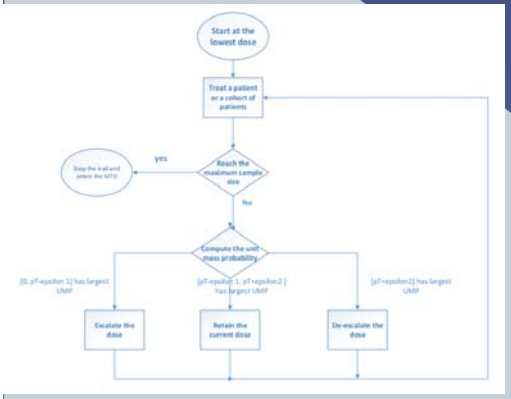
Physician specifies highest toxicity prob acceptable to treat future patients without dose de-escalation → $p_T + \varepsilon_2$

- Divides space for toxicity prob i: $[0, p_T - \varepsilon_1]$, $[p_T - \varepsilon_1, p_T + \varepsilon_2]$, $[p_T + \varepsilon_2, 1]$
- Make decision based on which interval has the largest unit probability mass (UPM) , defined as ratio of posterior density falling within the interval and interval length:
  1. $[0, p_T - \varepsilon_1]$ has the largest UPM: de-escalate
  2. $[p_T - \varepsilon_1, p_T + \varepsilon_2]$ has the largest UPM: retain current dose
  3. $[p_T + \varepsilon_2, 1]$ has the largest UMP: escalate

mTPI – Modified Toxicity Probability Interval Design

Ji et al (Clinial Trials 2010, JCO 2013)

Comparative Simulation Results with matched sample sizes



- 3+3 design has higher risks of exposing pts to toxic doses above MTD
- Compared to mTPI, 3+3 does not yield higher probs in identifying correct MTD
- mTPI is equally transparent as 3+3, costless to implement with free software, more flexible in practical situations

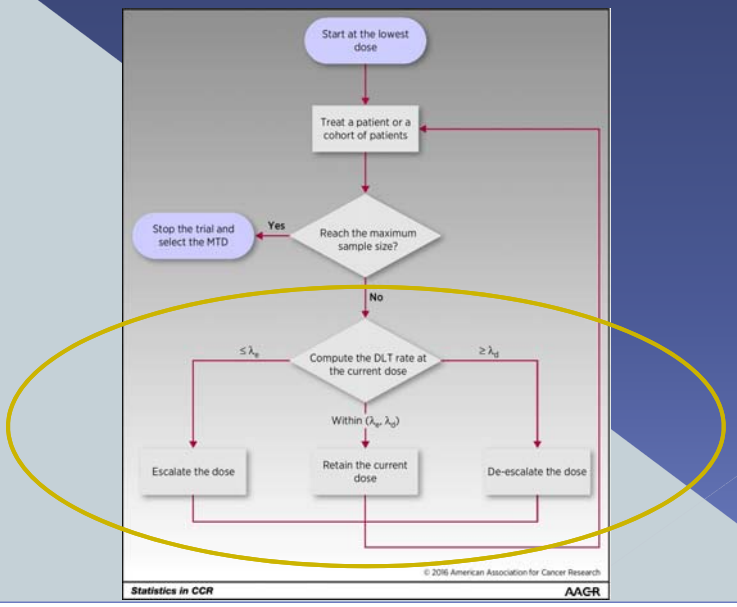
Bayesian Optimal Interval (BOIN) Design

Biostatistics Faculty: Suyu Liu, Ying Yuan, Ken Hess

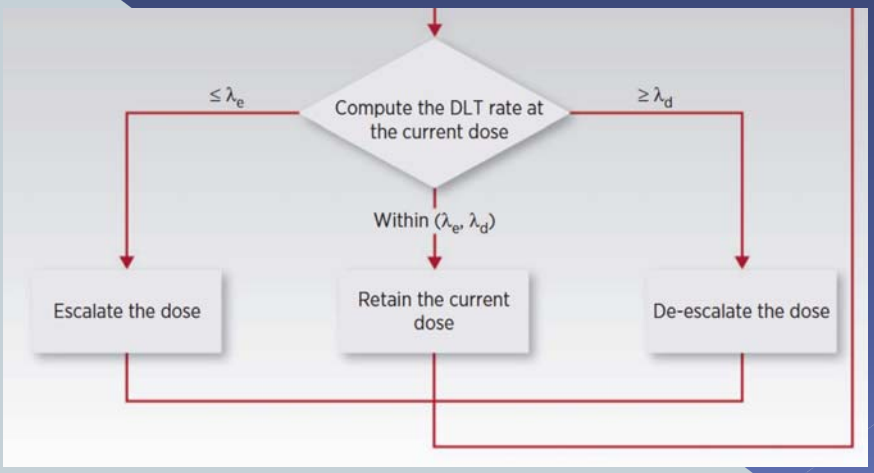
- BOIN design combines the simplicity of the 3+3 design with the superior performance of more complicated, model-based designs, e.g., continuous reassessment method (CRM).
- BOIN makes the decision of dose escalation and de-escalation simply by comparing the observed DLT rate at the current dose with a pair of fixed, predetermined dose escalation and de-escalation boundaries.
- Has been used in >20 trials in MDA and NCI.

Liu S and Yuan Y (2015) JRSS-C, 64, 507-523.
Yuan Y, Hess K, Hilsenbeck, S & Gilbert M (2016), *Clinical Cancer Research*, 22, 4291-4301.

Bayesian Optimal Interval (BOIN) Design



Bayesian Optimal Interval (BOIN) Design

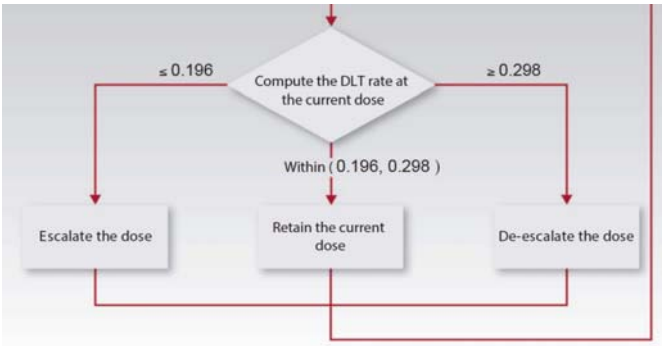


$$\text{DLT rate at the current dose} = \frac{\text{No. of patients experienced DLT at the current dose}}{\text{No. of patients treated at the current dose}}$$

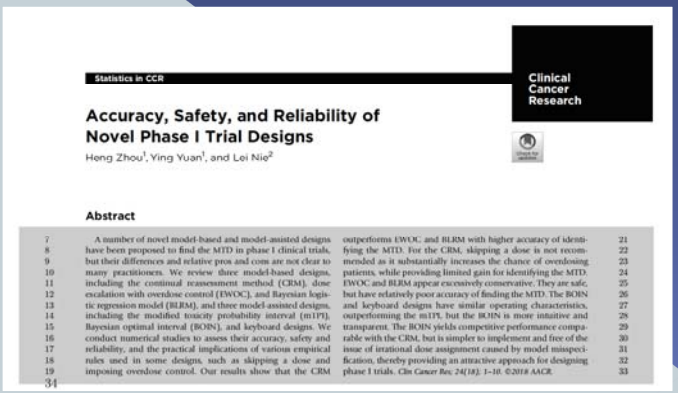
Escalation/Deescalation Boundaries

Table 1. Dose escalation and de-escalation boundaries

	Target toxicity rate for the MTD						
Boundary	0.1	0.15	0.2	**0.25**	0.3	0.35	0.4
λ_e (escalation)	0.078	0.118	0.157	**0.197**	0.236	0.276	0.316
λ_d (de-escalation)	0.119	0.179	0.238	**0.298**	0.358	0.419	0.479



- Although extremely simple, BOIN has very competitive performance.



Statistics in CCR

Clinical Cancer Research

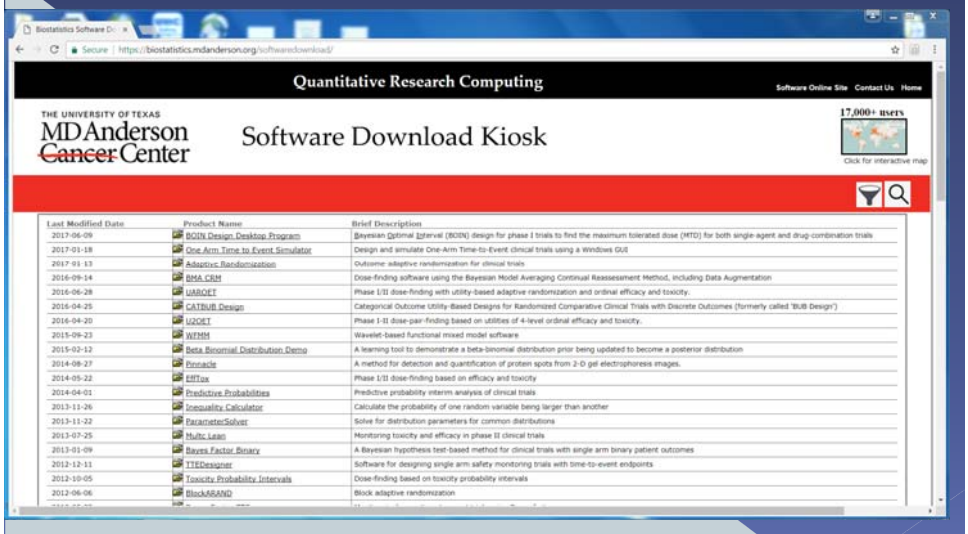
Accuracy, Safety, and Reliability of Novel Phase I Trial Designs

Heng Zhou[1], Ying Yuan[1], and Lei Nie[2]

Abstract

A number of novel model-based and model-assisted designs have been proposed to find the MTD in phase I clinical trials, but their differences and relative pros and cons are not clear to many practitioners. We review three model-based designs, including the continual reassessment method (CRM), dose escalation with overdose control (EWOC), and Bayesian logistic regression model (BLRM), and three model-assisted designs, including the modified toxicity probability interval (mTPI), Bayesian optimal interval (BOIN), and keyboard designs. We conduct numerical studies to assess their accuracy, safety and reliability, and the practical implications of various empirical rules used in some designs, such as skipping a dose and imposing overdose control. Our results show that the CRM outperforms EWOC and BLRM with higher accuracy of identifying the MTD. For the CRM, skipping a dose is not recommended as it substantially increases the chance of overdosing patients, while providing limited gain for identifying the MTD. EWOC and BLRM appear excessively conservative. They are safe, but have relatively poor accuracy of finding the MTD. The BOIN and keyboard designs have similar operating characteristics, outperforming the mTPI, but the BOIN is more intuitive and transparent. The BOIN yields competitive performance comparable with the CRM, but is simpler to implement and free of the issue of irrational dose assignment caused by model misspecification, thereby providing an attractive approach for designing phase I trials. *Clin Cancer Res; 24(18); 1–10. ©2018 AACR.*

Zhou H, Yuan Y, Hess K and Lei Nie (2018), *Clinical Cancer Research*, DOI: 10.1158/1078-0432.CCR-18-0168



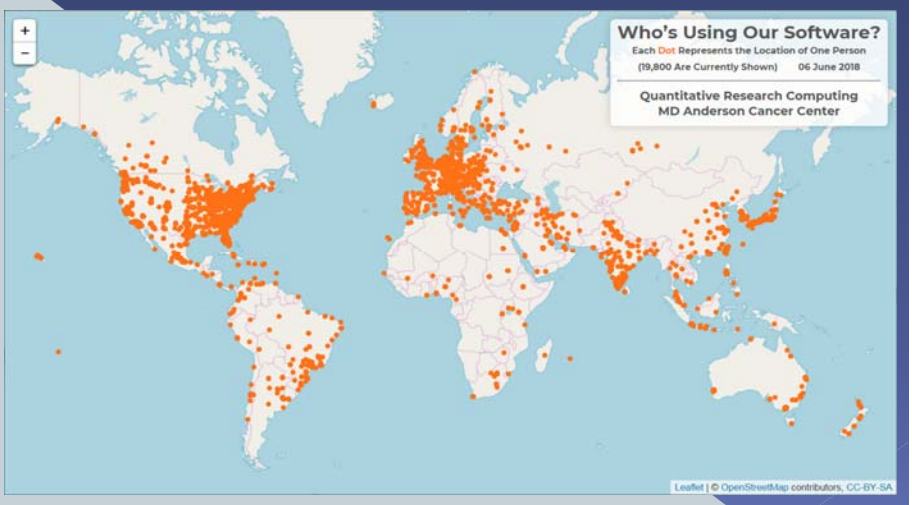
Quantitative Research Computing

THE UNIVERSITY OF TEXAS MD Anderson Cancer Center

Software Download Kiosk

17,000+ users

Last Modified Date	Product Name	Brief Description
2017-06-09	BOIN Design Desktop Program	Bayesian Optimal Interval (BOIN) design for phase I trials to find the maximum tolerated dose (MTD) for both single-agent and drug-combination trials
2017-01-18	One Arm Time to Event Simulator	Design and simulate One-Arm Time-to-Event clinical trials using a Windows GUI
2017-01-13	Adaptive Randomization	Outcome adaptive randomization for clinical trials
2016-09-14	BMA CRM	Dose-finding software using the Bayesian Model Averaging Continual Reassessment Method, including Data Augmentation
2016-06-28	UAROET	Phase I/II dose-finding with utility-based adaptive randomization and ordinal efficacy and toxicity.
2016-04-25	CATBUB Design	Categorical Outcome Utility-Based Designs for Randomized Comparative Clinical Trials with Discrete Outcomes (formerly called 'BUB Design')
2016-04-20	UZOET	Phase I-II dose-pair-finding based on utilities of 4-level ordinal efficacy and toxicity.
2015-09-23	WTMM	Wavelet-based functional mixed model software
2015-02-12	Beta Binomial Distribution Demo	A learning tool to demonstrate a beta-binomial distribution prior being updated to become a posterior distribution
2014-08-27	Pinnacle	A method for detection and quantification of protein spots from 2-D gel electrophoresis images.
2014-05-22	EffTox	Phase I/II dose-finding based on efficacy and toxicity
2014-04-01	Predictive Probabilities	Predictive probability interim analysis of clinical trials
2013-11-26	Inequality Calculator	Calculate the probability of one random variable being larger than another
2013-11-22	ParameterSolver	Solve for distribution parameters for common distributions
2013-07-25	Multc Lean	Monitoring toxicity and efficacy in phase II clinical trials
2013-01-09	Bayes Factor Binary	A Bayesian hypothesis test-based method for clinical trials with single arm binary patient outcomes
2012-12-11	TTEDesigner	Software for designing single arm safety monitoring trials with time-to-event endpoints
2012-10-05	Toxicity Probability Intervals	Dose-finding based on toxicity probability intervals
2012-06-06	BlockARAND	Block adaptive randomization



A semi-mechanistic dose-finding design in oncology using pharmacokinetic/pharmacodynamics modeling

Biostatistics faculty: Yisheng Li & Kim-Anh Do
Biostatistics postdoc fellow: Xiao Su

- Motivation
- Feasibility
- Performance
- Conclusion

Need for more efficient early phase trial designs

- High failure rates in phase III cancer clinical trials
- Known and potential reasons:
  - Narrow therapeutic index of cancer treatments
  - Small sample sizes of early phase trials
  - Limitations of existing dose-response models:
    1. Empirically based
    2. Difficulty in modeling schedule effects
    3. Difficulty in modeling effects of method of administration, including drug formulation, route of administration, and drug delivery system
- Essence of the limitations: ignores the mechanism of the drug effects that are often (partially) available
- A potential solution: mechanism-based dose-response modeling

A motivating Phase 1 trial at MD Anderson

Patients: with metastatic or locally advanced solid tumors
Study drug: a γ-secretase inhibitor
Mechanism of the drug effect in vivo: blocks Notch signaling via γ-secretase inhibition and produces a slower growing differentiation phenotype in human cancer cells

Treatment cycle: 10 days
Five total doses per cycle: 10mg, 14mg, 20mg, 23mg, 31mg
Method of administration: intravenous (IV)
Four schedules: 3-, 4-, 5-, and 6-dosing regimen

Primary goals: characterize the DLTs, overall safety, and identify an MTD for each schedule
Secondary objectives: characterize PK profiles
Target toxicity: 30%

PK/PD modeling and relationship with clinical outcome

PK model:

- dynamic dose(schedule) - concentration relationship
- characterizes what the body does to the drug

PD model:

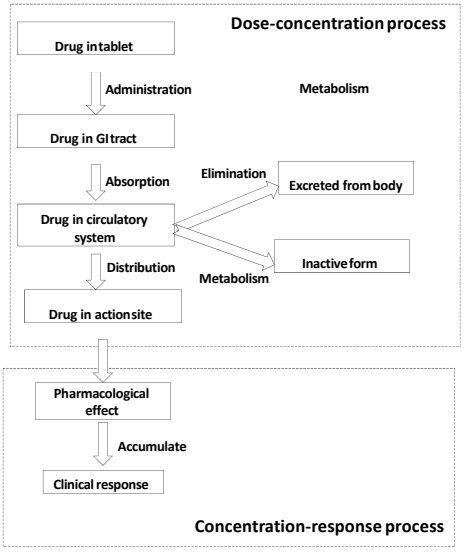
- dynamic concentration-effect relationship
- characterizes what the drug does to the body

Model the full PK/PD profiles, and model its effect on the DLT prob

Proposed joint model represents a Dose Concentration Effect Intensity (Clinical) Outcome (DCEO) framework

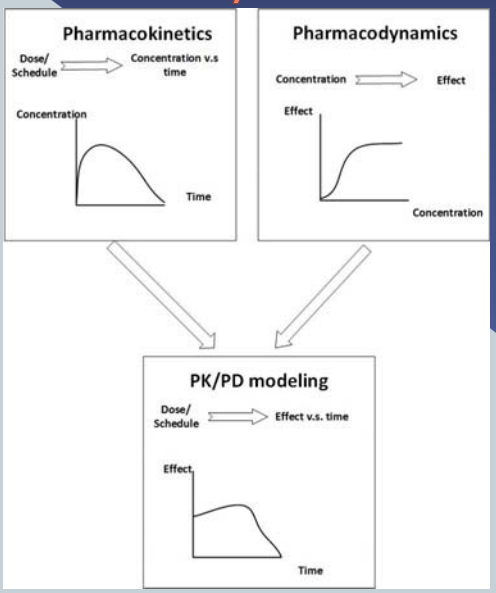
Proposed design is a Semi-Mechanistic Dose-Finding (SMDF) design

Mechanism



Motivation: Mechanistic dose-response modeling

PK/PD modeling



PD marker is latent

To identify PD model para's need to link PD marker to obs DLT outcome

PK/PD modeling links 2 important fields of pharmacology together, result in integrated PK/PD models describing dynamic drug effect intensity over time in response to a dosing regimen

Final step: link drug effect to the DLT outcome by a link function
$h(\eta) = 1 - \exp(-\eta)$
where η is cumulative effect of drug for patient over time interval $[0,t]$

PK/PD profiles based on preclinical studies

- PK/PD modeling/profile information based on PK/PD data from preclinical studies
- Uncertainty quantification of the PK/PD profile information base on PK/PD data from *in vitro* and *in vivo* studies

PK/PD data in clinical trials

- PK data almost always available in dose-finding trials
- PD data often available for efficacy evaluation in phase I/II or phase II studies

Some notation

Drug administration time: $s = (s_1, s_2, \ldots, s_K)$
Drug administration dosage: $d = (d_1, d_2, \ldots, d_K)$
Dose of one administration: d
Treatment regimen: $\tau = (d, s)$
Drug concentration in plasma at time t: $c(t)$
Drug concentration in gastrointestinal tract: $a(t)$ (when applicable)
Volume of distribution of the drug in the body: V
First-order elimination rate in one-compartment model: k_e
First-order absorption rate in one-compartment model: k_a
(when applicable)

Some initial conditions, e.g., $c(0) = d/V$

PK models

- IV administration:
  - Drug concentration profile for a one-compartment model with IV administration:

$$dc(t)/dt = -k_e c(t),\ c(0) = d/V$$

  - Closed-form solution:

$$c(t) = \frac{d}{V} e^{-k_e t}$$

  - Based on the superposition principle (Wang and Ouyang, 1998), we have:

$$c(t) = \sum_{k=1}^{K} \frac{d_k}{V} e^{-k_e(t-s_k)} I(t \geq s_k)$$

- Extravenous administration (EV): similarly developable based on solving $c(t)$ (and $a(t)$) analytically from a system of two differential equations

PD model

One of the most commonly used PD model under steady-state conditions is the following sigmoid E_{max}-model:

$$e(t \mid \phi, \psi) = \frac{E_{max} \times c(t \mid \phi)^{\gamma}}{EC_{50}^{\gamma} + c(t \mid \phi)^{\gamma}},$$

where $e(t \mid \phi, \psi)$ is the measured (or latent) effect, and $\psi = (E_{max}, EC_{50}, \gamma)$ are PD parameters, in particular, with EC_{50} being the concentration that causes 50% of the maximum effect E_{max}.

Additional notation

- A prespecified window for evaluating DLT: $[0, t^{ref}]$
- Drug cumulative effect up to time t:
  $\eta(t \mid \tau, \phi, \psi) = \int_0^t e\,[c(u \mid \tau, \phi); \psi]\ du$
- For one-compartment model with IV administration:

$$\eta(t|\tau, \theta) = \int_0^t \frac{Emax \times \left(\sum_{j=1}^{K} \frac{d_j}{V_c} e^{-k_{10}(u-s_k)} I(u \geq s_k) \right)^{\gamma}}{ED_{50}^{\gamma} + \left(\sum_{j=1}^{K} \frac{d_j}{V_c} e^{-k_{10}(u-s_k)} I(u \geq s_k) \right)^{\gamma}} du,$$

  which has an analytical form;
- Toxicity probability within the window $[0, t^{ref}]$ for patient i:

$$\pi_{\tau_i, \theta} = 1 - \exp\left[-\eta(t^{ref} \mid \tau_i, \phi, \psi)\right]$$

- Recorded concentrations at times t_{ij}: X_{ij}
- Clinical outcome (DLT) for patient i: $Y_i \in \{0, 1\}$

Models for PK data and clinical outcome

- Sampling model:

$$X_{ij} = c(t_{ij} \mid \boldsymbol{\tau_i}\phi) + \epsilon_{ij}, \text{ and } p(Y_i = 1) = \pi_{\boldsymbol{\tau_i}},$$

  with normal residuals $\epsilon_{ij} \sim Normal(0, \sigma^2)$.
- Independent prior distributions:
  - Unif$(0, 10)$ for k_e
  - Unif$(0, 5)$ for γ
  - Unif$(0, 100)$ for each of V, E_{max}, EC_{50}, and σ^2

Posterior inference

Posterior distribution:

$$\begin{aligned} p(\boldsymbol{\theta}|\mathcal{D}) &\propto L(\boldsymbol{X}, \boldsymbol{Y}|\boldsymbol{\theta}, \boldsymbol{\tau_i})g(\boldsymbol{\theta}) \\ &= L(\boldsymbol{Y}|\boldsymbol{\theta})L(\boldsymbol{X}|\boldsymbol{\theta})g(\boldsymbol{\theta}) \\ &= \prod_i \prod_j \frac{1}{\sigma\sqrt{2\pi}} exp\left\{-\frac{\left[x_{ij} - c(t_{ij}|\boldsymbol{\tau_i}, \boldsymbol{\theta})\right]^2}{2\sigma^2}\right\} \\ &\quad \times \prod_i \left(1 - \exp\left\{-\int_0^{t^{ref}} e\left[c(t)t_{ij}|\boldsymbol{\tau_i}, \boldsymbol{\theta}\right] du\right\}\right)^{Y_i} \\ &\quad \left(\exp\left\{-\int_0^{t^{ref}} e\left[c(t)t_{ij}|\boldsymbol{\tau_i}, \boldsymbol{\theta}\right] du\right\}\right)^{1-Y_i} \\ &\quad \times g(\boldsymbol{\theta}), \end{aligned}$$

where $\boldsymbol{\theta}$ denotes all unknown parameters, and $g(\boldsymbol{\theta})$ is the joint prior distribution for $\boldsymbol{\theta}$.

Posterior MCMC simulation using hyperrectangle slice sampling

Dose-Finding Design- Decision Rules

- Optimal regimen: defined as (15) $\tau^{opt} = argmin_{\tau} |\hat{p}(\tau|D) - \phi|,$
  Φ is the target toxicity probability
- Posterior updating: At any decision point, the posterior distribution of the unknown parameters is updated using all accumulated data at that point.
- Safety rules: To avoid overdosing, define regimen τ to be safe if

$$Pr\,[\hat{p}(\tau|D) \leq \phi] > \psi,$$

tuning para ψ

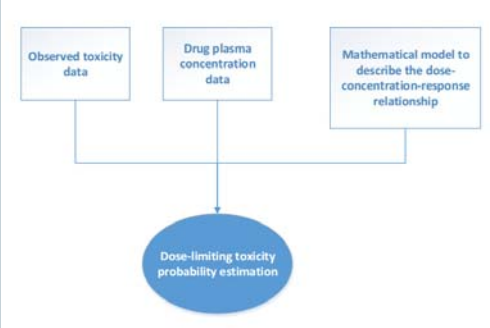
Algorithm 1 Semi-mechanistic dose finding (SMDF) design

Input: tuning parameters ϕ, ψ, n_{max}, c; hyperprior $g(\theta_{PD}, \sigma^2, \theta_{POP})$
Output: MTD τ^{opt}; estimated dose-toxicity curve (optional);

Allocate the first cohort of patients at the starting dose (lowest dose);
repeat
  Update τ^{opt} using (15) based on the current data;
  Allocate the next cohort at τ^{opt}, subject to no skipping of untried doses. When no DLT has occurred in any patient at any dose, allocate the next cohort at the next higher dose;
  if there is no safe dose **then**
    terminate the trial inconclusively;
  end if
until $n = n_{max}$ or trial is terminated.

SMDF

General idea to estimate toxicity probability



The SMDF design utilizes both toxicity data and drug concentration data by a joint model to describe the dose-concentration-response relationship simultaneously.

Simulation

- Five does levels, target toxicity of 0.3
- Evaluate OC's under 2 different clinical settings:
  - Max sample size 30, cohort size 3
  - Max sample size 2, cohort size 2

SMDF design:

- Concentration measurement time: (1, 3, 5, 10, 12, 24) hours
- t^{ref} = 7 hours
- Drug is administered by IV - 1 dosing regimen in 7-day trtmt cycle
- Dose levels: total dosage administered within a trtmt cycle
- Working model: *IV* (for PK) + E_{max} (for PD) ; tuning para $\psi = 0.1$
- η: cumulative toxicity effect of drug for pt i over time interval t

$$\eta_i(t|\boldsymbol{\tau},\theta) = \int_0^t \frac{Emax \times \left(\frac{k_{a,i}d}{V_{c,i}}\left[e^{-k_{e,i}(u-s)}\right]\right)^{\gamma}}{ED_{50}^{\gamma} + \left(\frac{k_{a,i}d}{V_{c,i}}\left[e^{-k_{e,i}(u-s)}\right]\right)^{\gamma}} du.$$

Simulation (con't)

CRM design:

- Toxicity prob at level j: $p_j = a_j^{\exp(\alpha)}$ $\alpha \sim N(0,1.34)$
- Implemented safety constraint by Zohar & Chevret (2000)
- Chose 2 sets of (skeleton) prior probabilities by Lee and Cheung (2009), to investigate sensitivity of CRM design to choice of skeleton

  (0.123, 0.204, 0.300, 0.402, 0.501)

  (0.062, 0.160, 0.300, 0.453, 0.594)

mTPI (modified Toxicity Prob Interval) design:

assume $\varepsilon_1 = \varepsilon_2 = 0.05$

BOIN design:

- Set highest toxicity prob deemed subtherapeutic = 0.18
- Set lowest toxicity prob deemed overly toxic = 0.42
- Cutoff to eliminate any overly toxic dose is 0.95

Simulation (con't)

PKPOP design:

- PK data collection scheme is the same as in the SMDF design
- Critical predictor AUC is estimated using the IV model

Two different methods to generate data:

1. Modeling framework of the SMDF design using IV + 5PL (5-para logistic model)

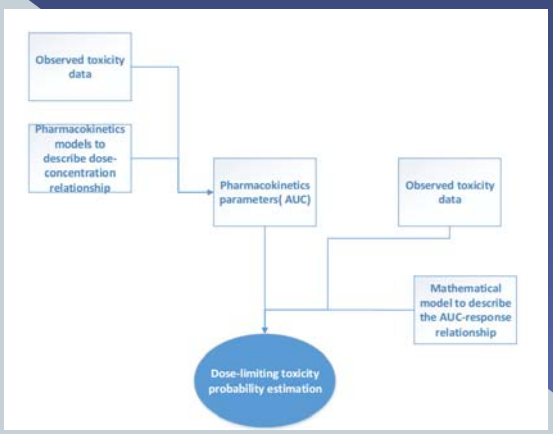
$$\eta_i(t|\boldsymbol{\tau},\boldsymbol{\theta}) = \int_0^t \frac{\beta_1}{\left[1+\left[\left(\frac{k_{a,i}d}{V_{d,i}}\left[e^{-k_{e,i}(u-\tau)}\right]\right)/\beta_3\right]^{\beta_4}\right]^{\beta_5}} du; \qquad (18)$$

2. Data generating framework in Ursino et al (2017), except using one compartment PK model IV, using a logit link of toxicity prob to AUC

This allows us to investigate the robustness of the performance of the SMDF design under more extreme scenarios

PKPOP

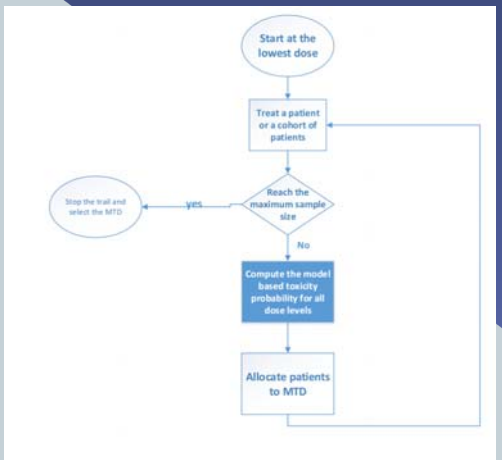
The general idea to estimate the toxicity probability



The PKPOP design utilizes both toxicity data and drug concentration data by a sequential procedure.

CRM, SMDF and PKPOP

The general allocation scheme of CRM, SMDF and PKPOP are quite similar



The major differences among the three designs are the method to compute toxicity probability(labeled in dark blue).

TABLE 1 Operating Characteristics of SMDF design, CRM designs using two initial guesses, BOIN design, mTPI design and PKPOP design. The data is simulated using IV+5PL model. The maximum sample size is set to be 30 and cohort size is 3.

		Dose(mg)	7	15	30	60	120			
		Toxicity	0.065	0.157	**0.261**	0.361	0.451	None	Tox	Npt
Scenario 1	SMDF	% Sel	0.01	0.10	**0.50**	0.31	0.078	0.00	0.26	29.95
		# Pts	3.69	5.91	11.03	6.87	2.44			
	CRM1	% Sel	0.00	0.10	**0.40**	0.38	0.11	0.00	0.26	30.00
		# Pts	4.24	5.85	9.38	7.59	2.95			
	CRM2	% Sel	0.00	0.13	**0.51**	0.31	0.05	0.00	0.25	30.00
		# Pts	3.91	6.42	11.59	6.37	1.71			
	BOIN	% Sel	0.01	0.18	**0.41**	0.30	0.10	0.00	0.24	29.97
		# Pts	4.32	8.20	9.75	5.68	2.02			
	mTPI	% Sel	0.01	0.17	**0.44**	0.29	0.08	0.00	0.24	29.97
		# Pts	4.17	8.19	10.33	5.46	1.82			
	PKPOP	% Sel	0.10	0.34	**0.34**	0.17	0.06	0.00	0.20	30.00
		# Pts	6.39	10.97	9.33	3.45	0.86			
		Toxicity	0.167	**0.301**	0.426	0.536	0.627	None	Tox	Npt
Scenario 2	SMDF	% Sel	0.16	**0.59**	0.22	0.01	0.00	0.02	0.30	29.66
		# Pts	8.32	13.40	6.73	1.05	0.15			
	CRM1	% Sel	0.19	**0.52**	0.26	0.02	0.00	0.01	0.30	29.79
		# Pts	9.56	11.47	6.92	1.59	0.26			
	CRM2	% Sel	0.18	**0.57**	0.22	0.02	0.00	0.01	0.30	29.71
		# Pts	9.78	12.26	6.46	1.12	0.10			
	BOIN	% Sel	0.23	**0.53**	0.18	0.03	0.00	0.02	0.29	29.58
		# Pts	10.84	12.37	5.09	1.15	0.13			
	mTPI	% Sel	0.24	**0.51**	0.20	0.03	0.00	0.01	0.28	29.61
		# Pts	10.54	12.94	5.21	0.87	0.05			
	PKPOP	% Sel	0.49	**0.38**	0.12	0.01	0.00	0.01	0.24	29.89
		# Pts	15.94	10.55	3.09	0.29	0.01			
		Toxicity	0.001	0.010	0.094	**0.301**	0.500	None	Tox	Npt
Scenario 3	SMDF	% Sel	0.00	0.00	0.08	**0.74**	0.18	0.00	0.24	30.00
		# Pts	3.00	3.02	4.99	13.11	5.89			
	CRM1	% Sel	0.00	0.00	0.12	**0.62**	0.27	0.00	0.26	30.00
		# Pts	3.01	3.07	4.70	11.29	7.93			
	CRM2	% Sel	0.00	0.00	0.09	**0.72**	0.19	0.00	0.25	30.00
		# Pts	3.00	3.02	4.96	13.12	5.90			
	BOIN	% Sel	0.00	0.00	0.17	**0.67**	0.16	0.00	0.22	30.00
		# Pts	3.00	3.23	7.50	11.41	4.85			
	mTPI	% Sel	0.00	0.00	0.17	**0.66**	0.17	0.00	0.22	30.00
		# Pts	3.00	3.18	6.74	12.33	4.74			
	PKPOP	% Sel	0.00	0.01	0.28	**0.60**	0.12	0.00	0.19	30.00
		# Pts	3.00	3.61	8.99	11.63	2.78			

%Sel: selection probability
Pts: average number of patients allocated at each dose
None: probability of inconclusive trials
Tox: average percentage of patients experiencing toxicity
Npt: average number of patients in the trial
CRM1: CRM design using initial guess (0.123, 0.204, 0.300, 0.402, 0.501)
CRM2: CRM design using initial guess (0.062, 0.160, 0.300, 0.453, 0.594)

Scenario1:
True MTD is third dose,
CRM2 selects true MTD with highest prob,
SMDF performance close to CRM2

Scenario 2 and 3:
True MTD is second and fourth dose levels,
SMDF selects true MTD with highest prob amongst 6 designs,
SMDF allocates most pts at true MTD and fewest pts to overly toxic dose levels

TABLE 1 (Continued)

		Toxicity	0.126	0.166	0.207	0.252	**0.298**	None	Tox	Npt
Scenario 4	SMDF	% Sel	0.01	0.11	0.28	0.29	**0.32**	0.00	0.21	29.89
		# Pts	4.81	6.32	8.23	5.71	4.82			
	CRM1	% Sel	0.02	0.08	0.26	0.29	**0.35**	0.00	0.21	29.87
		# Pts	5.75	5.36	6.93	6.10	5.73			
	CRM2	% Sel	0.02	0.14	0.33	0.29	**0.21**	0.20	0.20	29.92
		# Pts	5.77	6.79	8.42	5.61	3.33			
	BOIN	% Sel	0.04	0.13	0.21	0.28	**0.35**	0.01	0.20	29.83
		# Pts	5.80	7.14	6.79	5.46	4.65			
	mTPI	% Sel	0.05	0.15	0.23	0.28	**0.27**	0.01	0.20	29.79
		# Pts	6.32	7.55	6.78	5.24	3.88			
	PKPOP	% Sel	0.03	0.34	0.27	0.17	**0.20**	0.00	0.20	30.00
		# Pts	3.78	12.35	7.33	3.67	2.88			
		Toxicity	0.419	0.509	0.586	0.656	0.716	**None**	Tox	Npt
Scenario 5	SMDF	% Sel	0.32	0.04	0.00	0.00	0.00	**0.63**	0.44	18.80
		# Pts	15.09	3.20	0.48	0.03	0.00			
	CRM1	% Sel	0.49	0.03	0.00	0.00	0.00	**0.47**	0.43	22.59
		# Pts	19.47	2.28	0.69	0.15	0.00			
	CRM2	% Sel	0.53	0.02	0.00	0.00	0.00	**0.45**	0.44	21.90
		# Pts	20.01	1.71	0.15	0.03	0.00			
	BOIN	% Sel	0.40	0.05	0.00	0.00	0.00	**0.55**	0.43	20.23
		# Pts	16.61	3.21	0.42	0.00	0.00			
	mTPI	% Sel	0.41	0.07	0.00	0.00	0.00	**0.51**	0.37	14.7
		# Pts	10.96	3.35	0.38	0.00	0.00			
	PKPOP	% Sel	0.47	0.06	0.00	0.00	0.00	**0.47**	0.43	23.36
		# Pts	20.77	2.33	0.23	0.02	0.00			

%Sel: selection probability
Pts: average number of patients allocated at each dose
None: probability of inconclusive trials
Tox: average percentage of patients experiencing toxicity
Npt: average number of patients in the trial
CRM1: CRM design using initial guess (0.123, 0.204, 0.300, 0.402, 0.501)
CRM2: CRM design using initial guess (0.062, 0.160, 0.300, 0.453, 0.594)

Scenario 4:
True MTD is highest dose,
Toxicity of fourth dose level close to that of true MTD

Max SS=30:
CRM1 and BOIN select true MTD with highest prob,
SMDF is third best , selection prob close to CRM1/BOIN
Selection prob of CRM2, PKPOP designs much lower (approx. 0.2)

Max SS=20:
CRM1 select true MTD with highest prob,
SMDF is third best

Scenario 5:
All dose levels are toxic
A satisfactory design should recommend "inconclusive" with high prob

SMDF recommends "inconclusive" with high prob

Operating characteristics

TABLE 3 Operating Characteristics of SMDF design, CRM designs using two initial guesses, BOIN design, mTPI design and PKPOP design. The data is simulated using Ursino's model. The maximum sample size is set to be 30 and cohort size is 3.

		Dose(mg)	34.65	44.69	60.81	83.69	100.37			
		Toxicity	0.038	0.079	0.166	**0.303**	0.399	None	Tox	Npt
Scenario 1	SMDF	% Sel	0.00	0.01	0.22	**0.48**	0.29	0.00	0.24	29.91
		# Pts	3.45	3.56	6.99	8.68	7.23			
	CRM1	% Sel	0.00	0.01	0.17	**0.49**	0.33	0.00	0.24	30.00
		# Pts	3.51	3.60	6.42	9.45	7.02			
	CRM2	% Sel	0.00	0.01	0.22	**0.54**	0.23	0.00	0.22	30.00
		# Pts	3.47	3.89	8.16	9.53	4.96			
	BOIN	% Sel	0.00	0.02	0.25	**0.48**	0.25	0.00	0.21	30.00
		# Pts	3.53	4.73	8.24	8.72	4.77			
	mTPI	% Sel	0.00	0.02	0.26	**0.49**	0.23	0.00	0.21	30.00
		# Pts	3.55	4.82	8.15	9.12	4.35			
	PKPOP	% Sel	0.08	0.09	0.23	**0.33**	0.28	0.00	0.19	30.00
		# Pts	5.42	5.77	7.30	6.52	5.00			
		Toxicity	0.158	**0.261**	0.420	0.601	0.697	NA	Tox	Npt
Scenario 2	SMDF	% Sel	0.16	**0.53**	0.25	0.01	0.00	0.04	0.29	28.93
		# Pts	9.15	10.63	7.46	1.5	0.20			
	CRM1	% Sel	0.12	**0.53**	0.32	0.01	0.00	0.01	0.30	29.67
		# Pts	8.74	10.92	8.10	1.80	0.10			
	CRM2	% Sel	0.11	**0.57**	0.30	0.01	0.00	0.01	0.29	29.75
		# Pts	8.54	12.00	7.90	1.26	0.05			
	BOIN	% Sel	0.16	**0.51**	0.29	0.02	0.00	0.02	0.28	29.52
		# Pts	9.21	12.02	6.96	1.27	0.06			
	mTPI	% Sel	0.16	**0.50**	0.29	0.02	0.00	0.02	0.28	29.55
		# Pts	9.08	12.08	7.19	1.15	0.06			
	PKPOP	% Sel	0.30	**0.42**	0.24	0.02	0.01	0.01	0.26	29.68
		# Pts	11.85	10.96	5.78	0.87	0.13			
		Toxicity	0.419	0.563	0.725	0.854	0.906	**NA**	Tox	Npt
Scenario 3	SMDF	% Sel	0.37	0.01	0.00	0.00	0.00	**0.63**	0.44	18.32
		# Pts	15.80	2.03	0.46	0.02	0.00			
	CRM1	% Sel	0.52	0.02	0.00	0.00	0.00	**0.46**	0.44	22.34
		# Pts	19.68	2.23	0.41	0.02	0.00			
	CRM2	% Sel	0.54	0.02	0.00	0.00	0.00	**0.44**	0.44	23.05
		# Pts	20.24	2.38	0.41	0.02	0.00			
	BOIN	% Sel	0.42	0.02	0.00	0.00	0.00	**0.57**	0.44	20.06
		# Pts	17.14	2.71	0.21	0.00	0.00			
	mTPI	% Sel	0.46	0.03	0.00	0.00	0.00	**0.51**	0.38	14.82
		# Pts	12.02	2.62	0.17	0.00	0.00			
	PKPOP	% Sel	0.44	0.06	0.00	0.00	0.00	**0.50**	0.44	21.41
		# Pts	18.23	2.88	0.26	0.02	0.00			

%Sel: selection probability
Pts: average number of patients allocated at each dose
None: probability of inconclusive trials
Tox: average percentage of patients experiencing toxicity
Npt: average number of patients in the trial
CRM1: CRM designs using initial guess (0.123, 0.204, 0.300, 0.402, 0.501)
CRM2: CRM designs using initial guess (0.062, 0.160, 0.300, 0.453, 0.594)

Scenarios 1 and 2:
True MTD is second and fourth dose levels,
CRM2 selects true MTD with highest prob,

Performance of SMDF, CRM!, BOIN, mTPI designs are close to each other

Scenario 3:
All dose levels are toxic
An adequate design should recommend mostly "inconclusive" trials and terminate the trial as quickly as possible

SMDF recommends "inconclusive" correctly with high prob

TABLE 3 (Continued)

		Toxicity	0.019	0.042	0.099	0.202	**0.283**	NA	Tox	Npt
Scenario 4	SMDF	% Sel	0.00	0.00	0.02	0.23	**0.74**	0.00	0.19	29.97
		# Pts	3.16	3.08	4.01	6.34	13.39			
	CRM1	% Sel	0.00	0.00	0.02	0.19	**0.79**	0.00	0.19	30.00
		# Pts	3.17	3.11	4.02	6.57	13.13			
	CRM2	% Sel	0.00	0.00	0.04	0.31	**0.65**	0.00	0.18	30.00
		# Pts	3.16	3.21	4.84	8.18	10.61			
	BOIN	% Sel	0.00	0.00	0.04	0.31	**0.65**	0.00	0.17	30.00
		# Pts	3.17	3.64	5.33	7.96	9.91			
	mTPI	% Sel	0.00	0.00	0.05	0.34	**0.61**	0.00	0.17	30.00
		# Pts	3.18	3.70	5.22	8.34	9.56			
	PKPOP	% Sel	0.06	0.04	0.11	0.2	**0.60**	0.00	0.16	30.00
		# Pts	4.60	4.35	5.66	5.73	9.66			

%Sel: selection probability
Pts: average number of patients allocated at each dose
None: probability of inconclusive trials
Tox: average percentage of patients experiencing toxicity
Npt: average number of patients in the trial
CRM1: CRM designs using initial guess $(0.123, 0.204, 0.300, 0.402, 0.501)$
CRM2: CRM designs using initial guess $(0.062, 0.160, 0.300, 0.453, 0.594)$

Scenario 4 :
True MTD is of the highest dose level.
CRM1 selects true MTD with highest prob.

Performance of SMDF is close to that of CRM1

TABLE 4 Operating Characteristics of SMDF design, CRM designs using two initial guesses, BOIN design, mTPI design and PKPOP design. The data is simulated using Ursino's model. The maximum sample size is set to be 20 and cohort size is 2.

		Dose(mg)	34.65	44.69	60.81	83.69	100.37			
		Toxicity	0.038	0.079	0.166	**0.303**	0.399	None	Tox	Npt
Scenario 1	SMDF	% Sel	0.00	0.03	0.24	**0.39**	0.33	0.00	0.23	19.93
		# Pts	2.45	2.87	4.56	4.74	5.31			
	CRM1	% Sel	0.00	0.03	0.21	**0.38**	0.38	0.00	0.23	20.00
		# Pts	2.44	2.91	4.48	5.28	4.90			
	CRM2	% Sel	0.00	0.04	0.26	**0.45**	0.25	0.00	0.21	20.00
		# Pts	2.40	3.18	5.37	5.65	3.40			
	BOIN	% Sel	0.00	0.07	0.32	**0.39**	0.21	0.00	0.19	20.00
		# Pts	2.83	4.22	5.73	4.68	2.53			
	mTPI	% Sel	0.01	0.06	0.27	**0.34**	0.31	0.00	0.21	19.96
		# Pts	2.48	3.52	5.23	5.02	3.71			
	PKPOP	% Sel	0.06	0.14	0.21	**0.31**	0.31	0.00	0.20	19.98
		# Pts	3.24	3.91	4.65	4.22	3.97			
		Toxicity	0.158	**0.261**	0.420	0.601	0.697	NA	Tox	Npt
Scenario 2	SMDF	% Sel	0.20	**0.44**	0.28	0.04	0.00	0.04	0.30	19.31
		# Pts	6.42	6.24	4.82	1.47	0.38			
	CRM1	% Sel	0.18	**0.45**	0.31	0.04	0.00	0.02	0.30	19.69
		# Pts	6.60	6.55	4.67	1.55	0.32			
	CRM2	% Sel	0.16	**0.49**	0.31	0.03	0.00	0.01	0.29	19.88
		# Pts	6.31	7.47	4.77	1.21	0.13			
	BOIN	% Sel	0.21	**0.49**	0.25	0.0.	0.00	0.03	0.27	19.6
		# Pts	7.53	7.36	3.73	0.89	0.10			
	mTPI	% Sel	0.21	**0.40**	0.31	0.01	0.00	0.04	0.29	19.28
		# Pts	6.12	7.15	4.62	1.24	0.14			
	PKPOP	% Sel	0.25	**0.39**	0.28	0.04	0.00	0.04	0.29	19.41
		# Pts	6.65	7.05	4.28	1.19	0.23			
		Toxicity	0.419	0.563	0.725	0.854	0.906	**NA**	Tox	Npt
Scenario 3	SMDF	% Sel	0.41	0.04	0.00	0.00	0.00	**0.55**	0.45	13.13
		# Pts	10.45	2.09	0.51	0.08	0.00			
	CRM1	% Sel	0.54	0.05	0.00	0.00	0.00	**0.41**	0.44	15.22
		# Pts	12.42	2.27	0.47	0.05	0.01			
	CRM2	% Sel	0.60	0.06	0.00	0.00	0.00	**0.34**	0.44	16.49
		# Pts	13.47	2.59	0.40	0.03	0.00			
	BOIN	% Sel	0.51	0.04	0.00	0.00	0.00	**0.45**	0.44	14.76
		# Pts	12.35	2.11	0.28	0.02	0.00			
	mTPI	% Sel	0.47	0.07	0.00	0.00	0.00	**0.46**	0.40	10.82
		# Pts	7.78	2.63	0.38	0.03	0.00			
	PKPOP	% Sel	0.44	0.09	0.01	0.00	0.00	**0.47**	0.45	14.00
		# Pts	10.44	3.06	0.44	0.05	0.00			

%Sel: selection probability
Pts: average number of patients allocated at each dose
None: probability of inconclusive trials
Tox: average percentage of patients experiencing toxicity
Npt: average number of patients in the trial
CRM1: CRM designs using initial guess $(0.123, 0.204, 0.300, 0.402, 0.501)$
CRM2: CRM designs using initial guess $(0.062, 0.160, 0.300, 0.453, 0.594)$

TABLE 4 (Continued)

		Toxicity	0.019	0.042	0.099	0.202	**0.283**	None	Tox	Npt
Scenario 4	SMDF	% Sel	0.00	0.00	0.06	0.25	**0.69**	0.00	0.18	19.96
		# Pts	2.13	2.29	3.05	3.99	8.49			
	CRM1	% Sel	0.00	0.00	0.04	0.26	**0.70**	0.00	0.18	20.00
		# Pts	2.13	2.35	3.04	4.39	8.10			
	CRM2	% Sel	0.00	0.00	0.08	0.31	**0.61**	0.00	0.17	20.00
		# Pts	2.11	2.46	3.62	5.15	6.66			
	BOIN	% Sel	0.00	0.02	0.13	0.35	**0.50**	0.00	0.15	20.00
		# Pts	2.32	3.07	4.45	5.06	5.10			
	mTPI	% Sel	0.00	0.02	0.12	0.26	**0.60**	0.00	0.17	20.00
		# Pts	2.14	2.65	3.84	4.70	6.67			
	PKPOP	% Sel	0.04	0.05	0.10	0.20	**0.61**	0.00	0.16	20.00
		# Pts	2.67	2.97	3.62	3.88	6.85			

%Sel: selection probability
Pts: average number of patients allocated at each dose
None: probability of inconclusive trials
Tox: average percentage of patients experiencing toxicity
Npt: average number of patients in the trial
CRM1: CRM designs using initial guess $(0.123, 0.204, 0.300, 0.402, 0.501)$
CRM2: CRM designs using initial guess $(0.062, 0.160, 0.300, 0.453, 0.594)$

Simulation results summary

- SMDF design improves identification of MTD in most scenarios when true data generating process is similar, yet still different from , the IV+Emax model.
- Although CRM designs may perform best in some scenarios, their performance may be sensitive to the choice of the skeleton
- If the true data generating model is considerably different from the IV+Emax model, CRM designs using 2 initial guesses outperform the other 4 designs in scenarios 1 and 2, but perform worse in scenarios 3 and 4
- In scenarios where CRM performs best, SMDF performs second best
- Overall, SMDF design performs better than CRM, BOIN, mTPI, and PKPOP designs considering both the efficiency and robustness of the MTD allocation, as well as patient allocation.
- DCEO framework facilitates better learning of the dose-toxicity curve, which may facilitate better dose selection in subsequent trials

TABLE 5 Toxicity Probability Estimation

		Maximum Sample Size=30				Maximum Sample Size=20			
True Model	Scenario	SMDF		PKPOP		SMDF		PKPOP	
		Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE
IV+5PL	1	0.006	0.007	-0.017	0.025	0.017	0.010	-0.017	0.037
	2	0.021	0.013	-0.010	0.030	0.019	0.020	-0.005	0.046
	3	0.017	0.005	-0.001	0.014	0.025	0.006	0.000	0.022
	4	0.001	0.009	-0.047	0.021	0.015	0.015	-0.038	0.029
	5	0.069	0.026	0.016	0.020	0.069	0.035	0.002	0.035
Ursino	1	0.023	0.007	-0.017	0.016	0.025	0.010	-0.012	0.026
	2	0.034	0.019	0.007	0.040	0.034	0.028	0.026	0.067
	3	0.102	0.043	0.086	0.052	0.115	0.064	0.137	0.098
	4	0.022	0.005	-0.017	0.010	0.024	0.007	-0.013	0.016

Investigate performance of estimation of dose toxicity curves

Ursino et al (2017) demonstrated improved performance of estimation of dose-toxicity curves using their proposed designs , including the PKPOP design, over conventional designs

Table 5 compares bias and MSE of the estimated toxicity prob corresponding to tried dose levels using SMDF and PKPOP models

Conclusion: SMDF model estimated the toxicity probabilities with slightly lower MSE and comparable biases

5 | APPLICATION

We illustrate the SMDF design by applying it to the setting of the phase I trial of the γ-secretase inhibitor for metastatic or locally advanced solid tumors. We consider a hypothetical realization of the trial under the SMDF design, generating drug concentration data and outcomes using the IV + Emax model as shown in (17). The protocol allows for $Q = 5$ total dose levels of 7mg, 15mg, 30mg, 60mg and 120mg. The drug is administered by IV in a 28-day treatment cycle with a once-weekly treatment schedule. The primary objective of the study is to identify the MTD, defined as the dosage with the probability of DLT closest to the target $\phi = 0.3$. Characterization of the PK profile is a secondary objective. The plasma concentrations of the drug for all patients are to be evaluated at $s = (1, 3, 5, 10, 12, 24)$ hours after the first dosing. A maximum of $n = 30$ patients can be enrolled.

Figure 1 summarizes the allocation history of the hypothetical realization of the trial. The selected MTD is 15mg, which is the second dose level as the dash line indicated. A total of 12 patients are treated at the MTD, with 6 of these 30 patients experiencing DLT. The numbers above each point are the observed data for each new allocated cohort. The numbers before the slash is the number of patients experiencing DLT for the new cohort while the numbers after the slash is the number of patients allocated at the dose level for the new cohort.

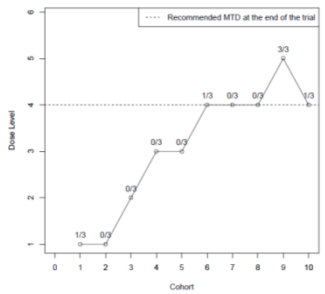


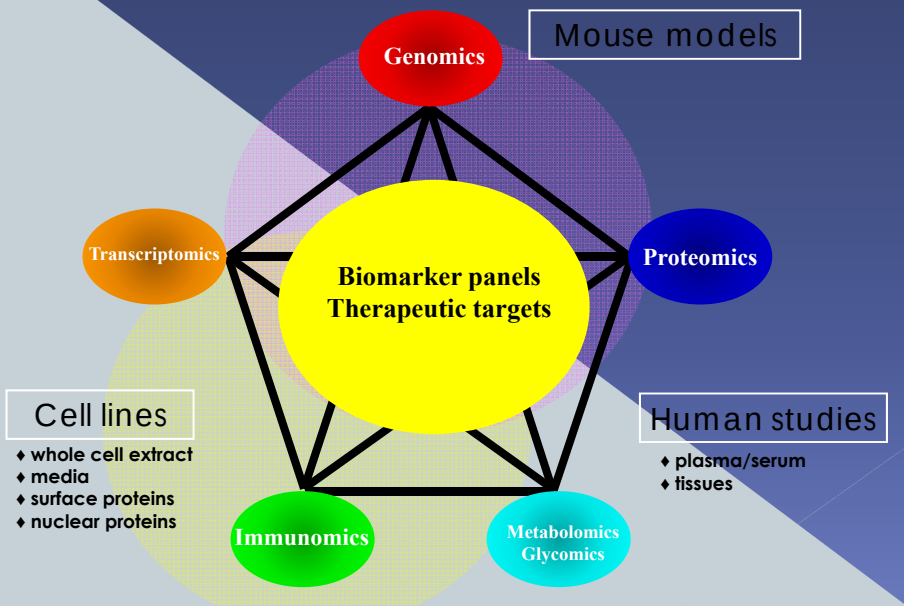
FIGURE 1 Allocation history of the hypothetical realization of the trial. The dash line indicated the recommended MTD at the end of the trial. The numbers above each point are the observed data for each new allocated cohort. The number before the slash is the number of patients experiencing DLT for the new cohort while the number after the slash is the number of patients allocated at this dose level for the new cohort.

Remarks

- We proposed a new modeling framework for the relationship for dose-finding designs in single- and multiple-schedule settings
- The proposed modeling framework and resulting designs look feasible and perform well, as compared to select common phase I designs
- The proposed modeling framework can be readily extended to a large number of early phase clinical trial settings
- Examples of some interesting issues in the extension of the designs:
  - Availability of analytic solutions to the system of differential equations that characterize the PK profiles, especially in population PK modeling
  - Applicability of superposition principle
  - Properties of the baseline hazard function when modeling time-to-event outcomes

Integration Perspective

Courtesy of S Hanash



Motivation to Integrate Different Data Platforms

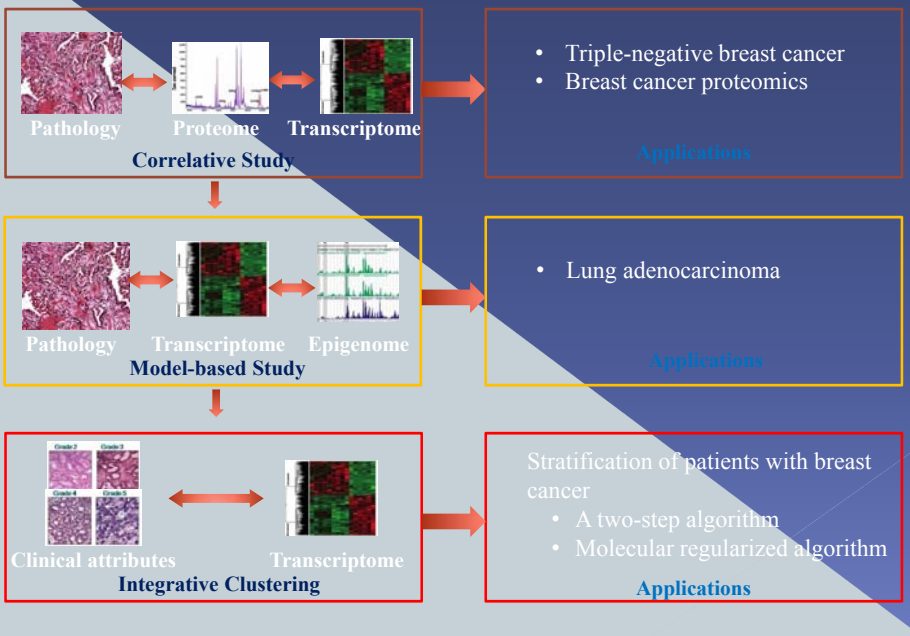
- The genome on its own "has turned out to be a relatively poor source of explanation for the differences between cells or between people" (Baines 2001)
- The clinical management of cancer and other complex diseases can be substantially improved by "integrative" approaches
- Distinct data types provide a different, only partly independent and complementary view on most complex, disease-related questions

- Biological complexity requires multi-dimensional and multi-system analysis
- Biological analysis requires an appreciation of context

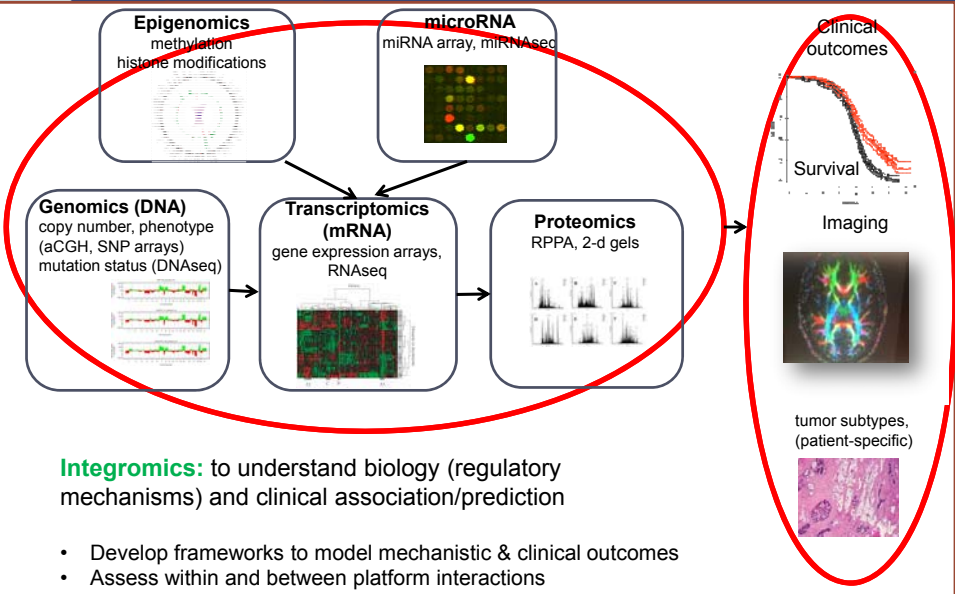
Understanding Cancer's Complexity

- Pathway analysis
- Disease modeling
- Cancer gene discovery
- Biomarkers
- Cancer Drug Development

From Correlation to Integration



Integromics (a biostatistician's perspective)



Integromics: to understand biology (regulatory mechanisms) and clinical association/prediction

- Develop frameworks to model mechanistic & clinical outcomes
- Assess within and between platform interactions
- Imaging + genomics — multiple levels of correlation

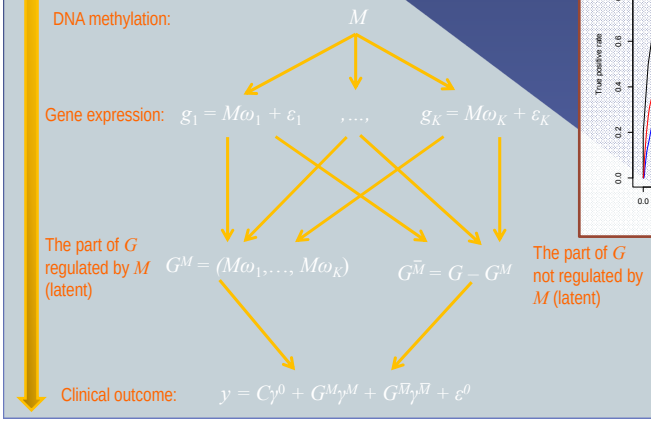
iBAG – Integrative Bayesian analysis of genomic data

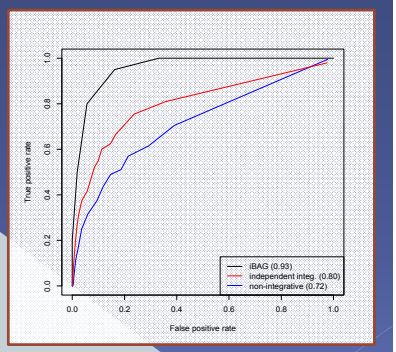
Bioinformatics, 2013

Application: application in glioblsastoma now extended to the colorectal cancer moon shot

TEAM: Wang, Balasundayuthapani, Morris, Do (Biostat); Broom, Manyam, (BCB); and glioblastoma researchers

Information flow

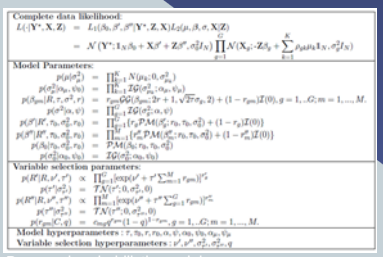




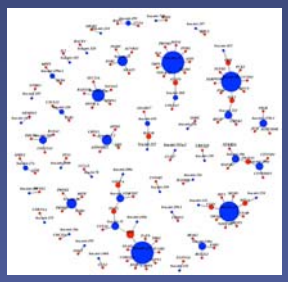
Integrating gene expression, miRNA with clinical outcome

Application: TCGA kidney cancer

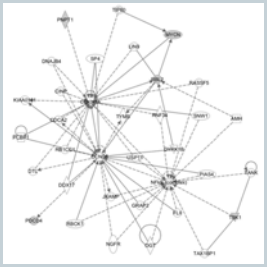
TEAM: Francesco Stingo, Thierry Chekouo, James Doecke, Kim-Anh Do (*Biometrics*, 2014)



Proposed probabilistic model



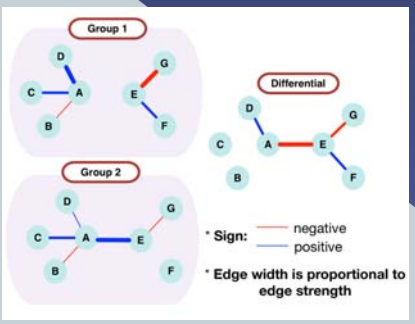
miRNA regulatory network using a threshold of 0.5 on its marginal posterior probability. The diameter of each node is proportional to its number of edges. Thicker edges correspond to higher marginal posterior probabilities. Red nodes correspond to genes; blue nodes correspond to miRNAs.



Gene regulatory network obtained from IPA. From the six genes identified, five (colored grey) could be mapped to a regulatory network based around NF B, CCND1 and CDKN1A, markers of immune response, cell cycle activity and tumor suppression.

DINGO: DIfferential Network analysis in GenOmics

TEAM: Min Jin Ha, Veera Baladandayuthapani, Kim-Anh Do (Bioinformatics, 2015)



DINGO: Model

- ``passenger" event : some conditional dependencies are shared across groups
- ``driver" event : other conditional dependencies are unique to the groups
- Decompose the conditional dependencies into a *global* *roup-*

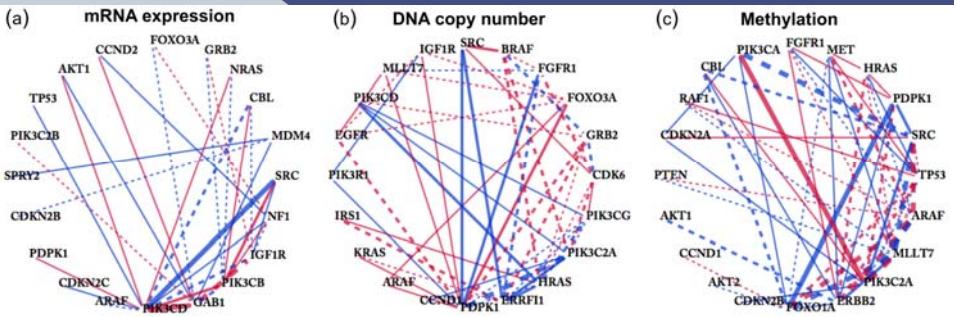
- We construct all $p(p-1)/2$ test statistics which can be constructed as a *differential score* as

$$\textit{Differential Score: } \delta_{ab}^{(12)} = \frac{\hat{\phi}_{ab}^{(1)} - \hat{\phi}_{ab}^{(2)}}{s_{ab}^{B}}. \quad (3)$$

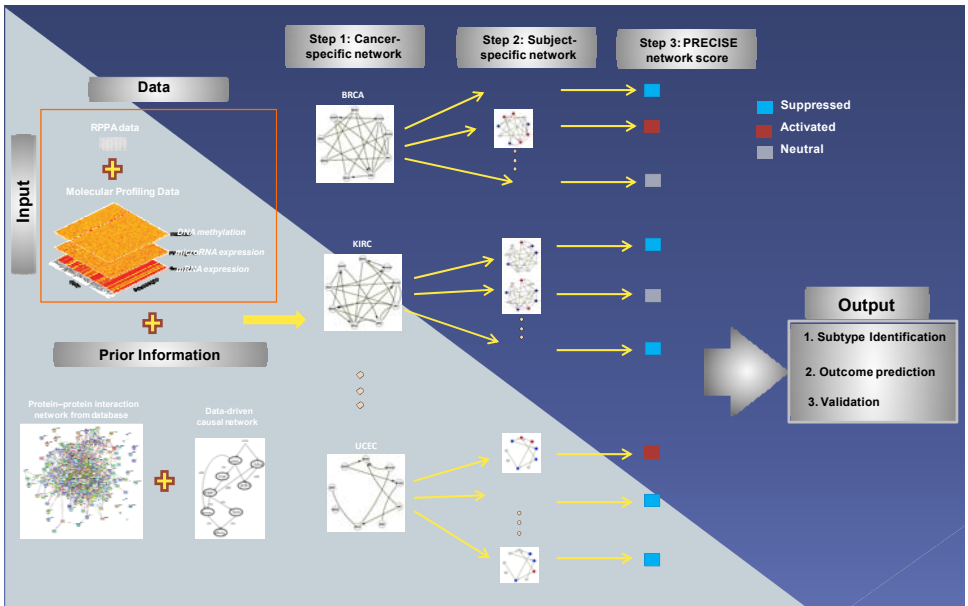
where $\hat{\phi}_{ab}^{(1)}$ and $\hat{\phi}_{ab}^{(2)}$ are Fisher's Z transformation of the estimates of $\rho_{ab}^{(1)}$ and $\rho_{ab}^{(2)}$, respectively, and s_{ab}^{B} is the bootstrap estimate of standard error.

DINGO: TCGA Glioblastoma

- The Cancer Genome Atlas (Glioblastoma) mRNA expression, DNA copy number and Methylation data
- 83 long-term survivors and 73 short-term survivors



- The blue (red) edges = positive (negative) differences
- The solid (dashed) lines = conserved (differential) signs
- The thickness of the edges = strength of the association



PRECISE: Personalized Cancer-Specific Integrated Network Estimation

TEAM: Baladandayuthapani, Ha, Do (Biostat); Mills (Systems Biology); Liang, Akbani (BCB)

Pathways

We focus on 12 pathways

Supplementary Table S2. Pathways and gene/protein names

	Pathway	Genes
1	Apoptosis	BAK1, BAX, BID, BCL2L11, CASP7, BAD, BCL2, BCL2L1, BIRC2
2	Breast reactive	CAV1, MYH11, RAB11A, RAB11B, CTNNB1, GAPDH, RBM15
3	Cell cycle	CDK1, CCNB1, CCNE1, CCNE2, CDKN1B, PCNA, FOXM1
4	Core reactive	CAV1, CTNNB1, RBM15, CDH1, CLDN7
5	DNA damage re	TP53BP1, ATM, BRCA2, CHEK1, CHEK2, XRCC5, MRE11A, TP53, RAD50, RAD51, XRCC1
6	EMT	FN1, CDH2, COL6A1, CLDN7, CDH1, CTNNB1, SERPINE1
7	PI3K/AKT	AKT1, AKT2, AKT3, GSK3A, GSK3B, CDKN1B, AKT1S1, TSC2, INPP4B, PTEN
8	RAS/MAPK	ARAF, JUN, RAF1, MAPK8, MAPK1, MAPK3, MAP2K1, MAPK14, RPS6KA1, YBX1
9	RTK	EGFR, ERBB2, ERBB3, SHC1, SRC
10	TSC/mTOR	EIF4EBP1, RPS6KB1, MTOR, RPS6, RB1
11	Hormone recep	ESR1, PGR, AR
12	Hormone signa	INPP4B, GATA3, BCL2

Rewiring protein signaling pathways in cancer

We investigate cancer-specific integrated networks across all cancer sites.

Aim 1: **Pathway activity**

The pathway activity is measured by the connectivity in the corresponding integrated network. If the pathway contains p number of proteins, the connectivity is defined by

$$\text{Connectivity score} = \frac{\text{number of edges between proteins}}{\text{total number of possible edges, } \binom{p}{2}},$$

where $0 \leq \text{Connectivity score} \leq 1$. **Higher connectivity score** means higher pathway activity (more cross signaling between proteins).

Aim 2: **Edge consistency**

We investigate tumor-specific/conserved edges. For an edge, the number of cancer sites that hold the edge are measured.

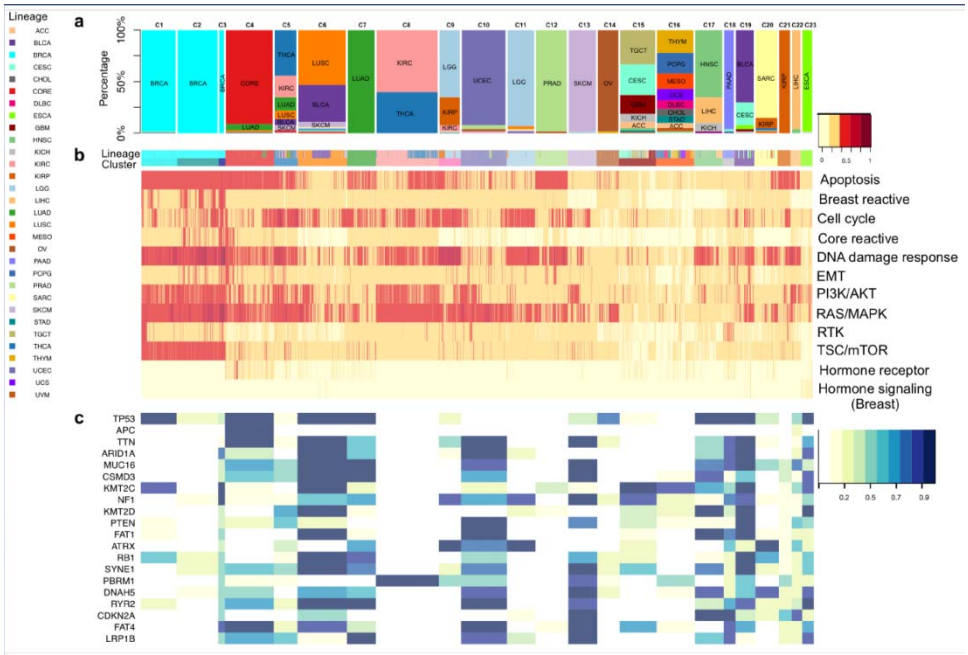


Figure 2.

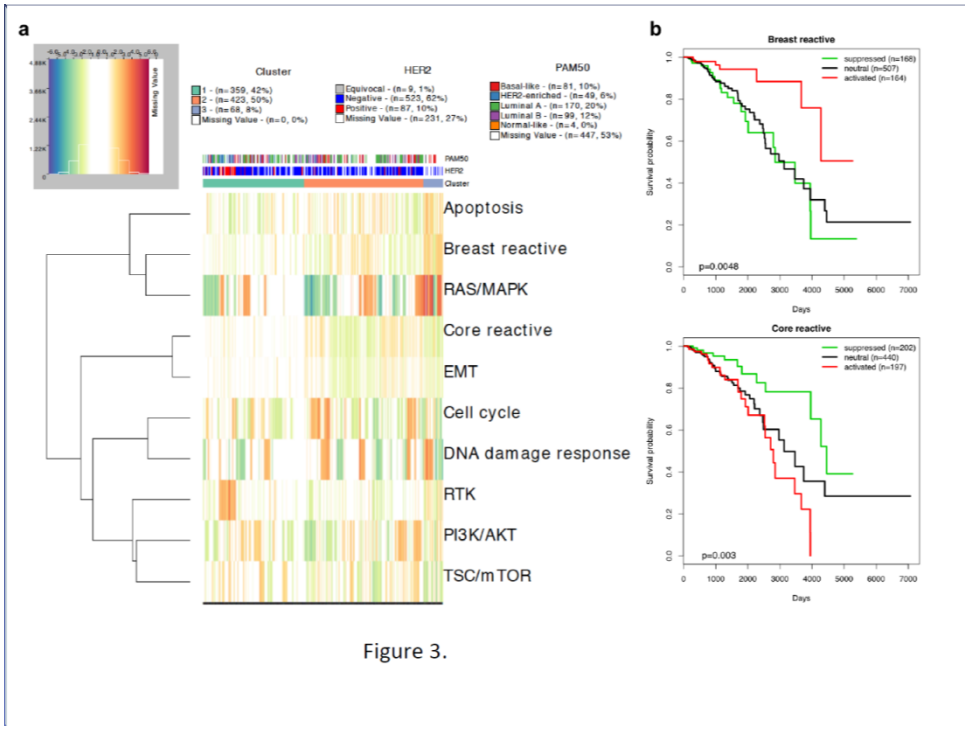


Figure 3.

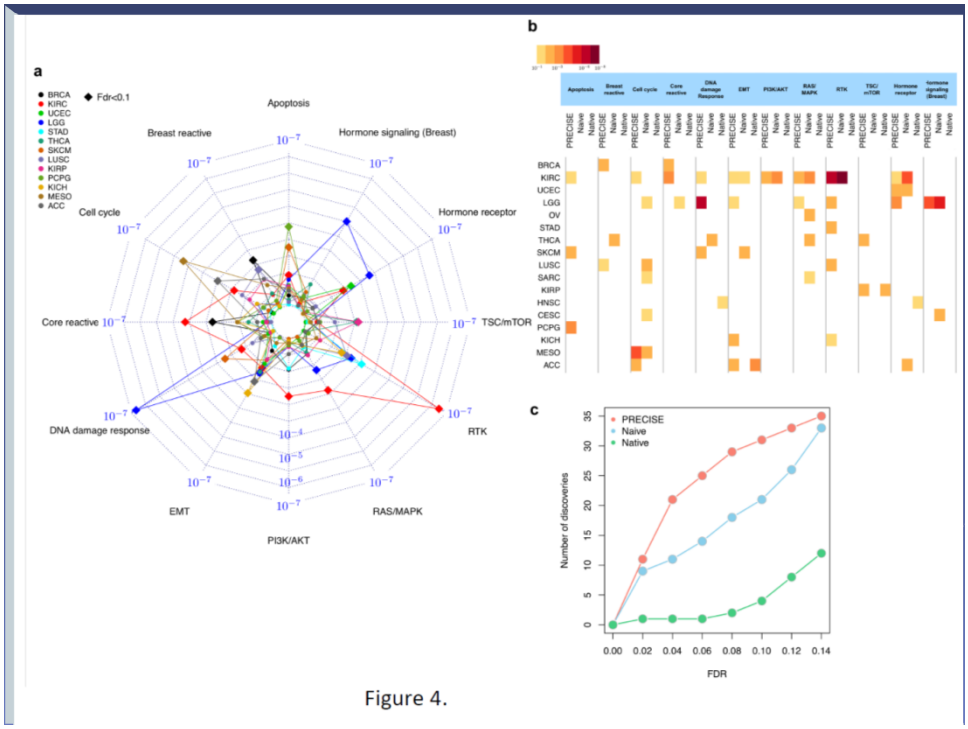


Figure 4.

Departmental goals for the next five years

- **Retain** and **recruit** faculty, researchers, and analysts to implement established and new impactful research areas (faculty recruitment committees consist of external members)
- Maintain and increase if possible the current competitive level of external research funding. Plan for a P01 in biostatistics methodology for impactful research areas
- Continue to engage with local institutions, including GSBS, Rice University, and Texas A&M University to strengthen our joint graduate programs, with emphasis on innovative courses and more effective recruitment processes to attract students with excellent quantitative skills
- The microbiome is an important factor in gene–environmental interactions that influence human disease susceptibility. Improve our methodology & collaborative research regarding the human microbiota as it relates to cancer development, antitumor immune responses, imaging modalities, and clinical efficacy of immunotherapies and other cancer treatments, and increase efforts in biomarker-driven clinical trial designs

"This could be the discovery of the century. Depending, of course, on how far down it goes."

If we can predict the move of our opponent, we can have a better chance of winning the game

Ingmar Bergman movie – The Seventh Seal