Design Concept for a Confirmatory Basket Trial

Robert A. Beckman, MD

[1]Professor of Oncology & of Biostatistics, Bioinformatics, and Biomathematics

Lombardi Comprehensive Cancer Center and Innovation Center for Biomedical Informatics

Georgetown University Medical Center

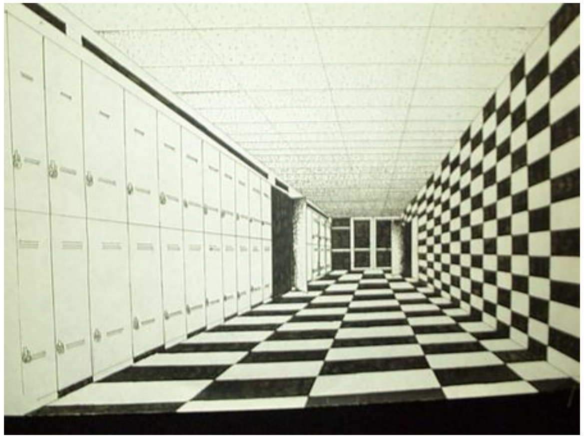
Founder and Chief Scientific Officer, Oncomind, LLC

Acknowledgements

- Co-authors on the present work:
  - **Cong Chen—led group; co-led concept development; led all statistical and simulation work**
  - Zoran Antonijevic, Amgen
  - Rasika Kalamegham, Genentech
- Pathway design subgroup, additional members:
  - Christine Gausse, Merck
  - Sebastian Jobjorrnsson, Chalmers
  - Lingyun Liu, Cytel
  - Sammy Yuan, Merck
  - Yi (Joey) Zhou, Ultragenyx
  - Advisor: Sue-Jane Wang, FDA
- Pathway design subgroup is one of 5 working subgroups of the **DIA Small Populations Workstream**, a group of 50 statisticians and clinicians from industry, academia, and national health authorities (FDA and EMEA)
- Small populations workstream is part of **DIA Adaptive Design Scientific Working Group (ADSWG),** a group of 200 statisticians and clinicians from industry, academia, and national health authorities (FDA and EMEA)

A Different Perspective

- Mathematical biology
- Cancer therapy development
  - 23 therapies first in man
  - 5 therapies advanced to late development
    - Her3 antagonist antibody
    - Pan-alpha integrin antibody
    - Anti IL6 antibody
    - Anti IGFR antibody
    - DR5 agonist antibody
  - 2 therapies approved:
    - Topotecan for small cell lung cancer
    - Bicalutamide for adjuvant therapy of prostate cancer
  - Small and large molecules targeting
    - Signal transduction
    - Repair
    - Angiogenesis
    - Developmental pathways
  - DNA vaccines, immunoliposomes, antibody-drug conjugates



Small Populations Within A Common Disease

- The increasing discovery of molecular subtypes of cancer leads to small subgroups that actually correspond to orphan or "niche" indications, even within larger tumor types
- Enrolling enough patients for confirmatory trials in these indications may be challenging.
- The shift to a molecular view of cancer requires a corresponding paradigm shift in drug development approaches
- Exclusive use of "one indication at a time" approaches will not be sustainable

A Major Challenge

- Cancer is becoming largely a collection of diseases defined by molecular subtype with low prevalence, even within major tumor types
- Enrolling enough subjects for confirmatory trials in these indications in a timely fashion is challenging
- Exclusive use of “one indication at a time” approaches will not be sustainable

Shift to molecular view of cancer → **Paradigm shift in drug development**

Rare Diseases

- Defined in Europe as prevalence of < 1 in 2K
  - Defined in US as prevalence of < 200K
- Up to 7000 rare diseases worldwide
- Up to 1 in 10 individuals worldwide affected
- Same issues with enrollment and difficult development as alluded to for biomarker defined subsets

Approaches to development based on predictive biomarkers

- Optimized co-development of a single drug and its companion diagnostic
  - Gives a clear hypothesis and answer and still has a role in selected instances
  - Will be challenging to do in niche indications
- "Umbrella" trials
  - One tumor type with multiple drugs and predictive biomarkers
  - Patients are matched to drugs based on predictive biomarkers
  - Cooperation among multiple sponsors
  - Examples: BATTLE, I-SPY, Lung-MAP
- "Basket" or "bucket" trials
  - Multiple tumor types with one drug and predictive biomarker
  - Approval based on pooled analysis
  - Premise is that molecular subtype is more fundamental than histology
  - Single sponsor

Agenda

- Introduction
- General Design Concept for a Basket Trial
- Challenges of Basket Trials and Recommendations for Overcoming Them
- Detailed Design Considerations
- Conclusions

The Original Basket: Imatinib B2225

186 patients with

40 different malignancies with known genomic mechanisms of activation of imatinib target kinases

KIT, PDGFRA or PDGFRB

Imatinib 400- 800 mg BID primary Endpoint ORR

Synovial Sarcoma	Aggressive Fibromatosis	Dermato-fibrosarcoma Protuberans	Aggressive Systemic mastocytosis	Hyper-eosinophilic syndrome	Myelo-proliferative disorder
1/16 (6%)	2/20 (10%)	10/12 (83%)	1/5 (20%)	6/14 (43%)	4/7 (58%)

13 centers in consortium: North America, Europe, Australia

Lead to supplemental indications for these 4 subsets after pooling with other trials and case reports

Blumenthal. Innovative trial designs to accelerate the availability of highly effective anti-cancer therapies: an FDA perspective, AACR 2014

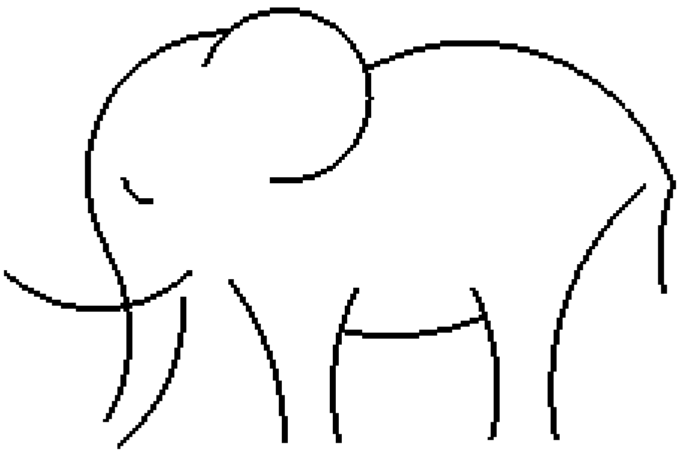
Basket Trials to Date

- A similar design to Imatinib B2225 was endorsed at a Brookings/Friends Conference in 2011
- Common features:
  - Exploratory and opportunistic in nature
  - Single-arm trials with ORR as primary endpoint
  - Intend to use pooled population for primary analysis to gain broader indication across tumor types (individual tumor type is not adequately powered)
  - Involve ***possibly*** transformative medicines in patients with great unmet need and ***seemingly*** exceptionally strong scientific rationale

Issues

- Clinical data to support pooling my be limited, and treatment effect may differ between tumor types
  - Vemurafenib works in melanoma with BRAF V600E mutation but not colorectal cancer with same mutation
- Not all drugs hoped to be transformational live up to this promise
- Response rate may not predict overall survival
- Single arm trials are subject to patient selection bias
- Predictive effect of a biomarker is confounded with the prognostic value which is often unknown
- Health authorities can be non-committal upfront

There Is An Elephant in the Room: Effective, but Not Transformational Drugs



Effective Drugs

- Pediatric acute lymphoblastic leukemia went from being a death sentence to ca 85% cure rate based on incremental advances using merely effective drugs.
- Is this merely a tale from the distant past?

Where Have Effective Drugs Gone?

- They far outnumber transformational drugs, and are in numerous pharmaceutical pipelines in great numbers, **BUT**
  - They have fewer press releases and press conferences
  - They are less likely to be published in *Lancet, Nature, New England Journal of Medicine, or Science*
  - They are less likely to be featured when a distinguished academic visits the company
- They may include:
  - Targeted small molecules
  - Antibodies and antibody drug conjugates
  - Compounds designed to improve further on the results of immunotherapy

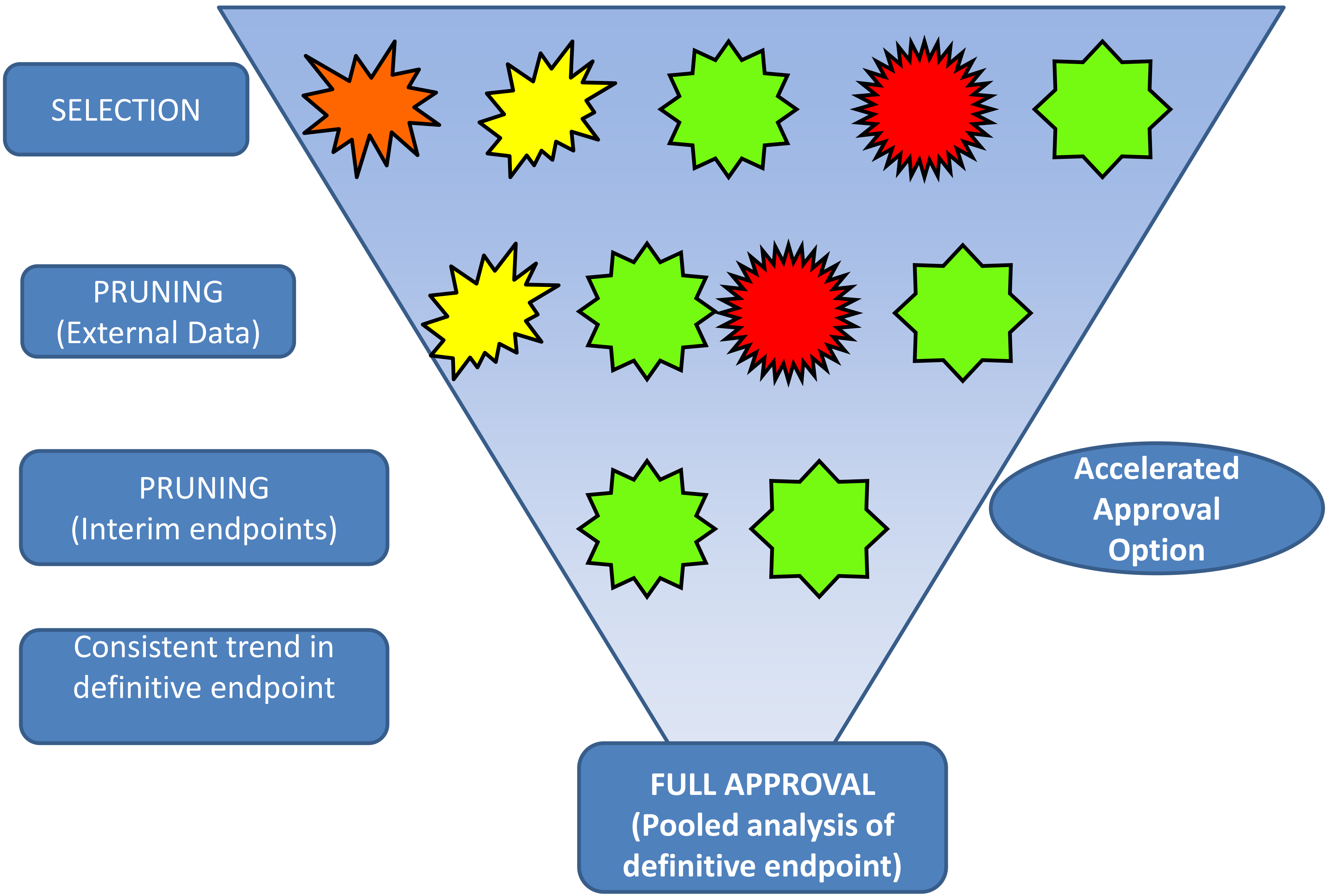
Dumb Tumors, Smart Tumors, and Low Hanging Fruit

Beckman, Robert A., Schemmann, Gunter S., and Yeang, Chen-Hsiang. Impact of genetic dynamics and single-cell heterogeneity on development of nonstandard personalized medicine strategies for cancer. Proceedings of the National Academy of Sciences USA, Published online before print August 13, 2012, 109: 14586-14591 (2012).

DIA Small Population Pathway Subteam

- Can we develop a **generalizable confirmatory basket design concept with statistical rigor?**
  - Applicable not only to exceptional cases, but to all effective medicines in any line of therapy
  - Follow existing accelerated and standard approval pathways to increase drug approvability
  - **Be complementary to current basket trial methods that are either exploratory, or confirmatory for transformational drugs only**
- This would have multiple benefits
  - Increase and accelerate access to effective medicines for patients in niche indications
  - Provide sponsors with cost-effective options for development in niche indications
  - Provide health authorities with more robust packages for evaluation of benefit and risk

GENERAL DESIGN CONCEPT


SELECTION
PRUNING
(External Data)
PRUNING
(Interim endpoints)
Consistent trend in
definitive endpoint
Accelerated
Approval
Option
FULL APPROVAL
(Pooled analysis of
definitive endpoint)

Features of the Design (I)

- Tumor histologies are grouped together, each with their own control group (shared control group if common SOC)
- Randomized control is preferred
  - Single arm cohorts with registry controls may be permitted in exceptional circumstances as illustrated by imatinib B225 and others
- In an example of particular interest, each indication cohort is sized for accelerated approval based on a surrogate endpoint such as progression free survival (PFS)
  - This may typically be 25-30% of the size of a Phase 3 study
- Initial indications are carefully selected as one bad indication can spoil the entire pooled result

Features of the Design (II)

- Indications are further "pruned" if unlikely to succeed, based on:
  - External data (maturing definitive endpoint from Phase 2; other data from class); IN FUTURE: real world data based on off label use
  - Internal data on surrogate endpoint
- Sample size of remaining indications may be adjusted based on pruning
- Type I error threshold will be adjusted to control type I error (false positive rate) in the face of pruning
  - Pruning based on **external** data does not incur a statistical penalty
  - Discussed in more detail later in talk
- Study is positive if pooled analysis of remaining indications is positive for the primary definitive endpoint
  - Remaining indications are eligible for full approval in the event of a positive study
  - Some of the remaining indications may not be approved if they do not show a trend for positive risk benefit as judged by definitive endpoint

CHALLENGES OF BASKET DESIGNS AND RECOMMENDATIONS FOR OVERCOMING THEM

Challenge 1: Risks of Pooling

- One of more bad indications can lead to a failed study for all indications in a basket
- Histology can affect the validity of a molecular predictive hypothesis, in ways which cannot always be predicted in advance
  - Vemurafenib is effective for BRAF V600E mutant melanoma, but not for analogous colorectal cancer (CRC) tumors
  - This was not predicted in advance but subsequently feedback loops leading to resistance were characterized

Addressing challenge 1

- Basket trials are recommended primarily after there has been a lead indication approved (by optimized conventional methods) which has validated the drug, the predictive biomarker hypothesis, and the companion diagnostic
  - Example, melanoma was lead indication preceding Brookings trial proposal in V600E mutant tumors
- Indications should be carefully selected
- Indications should be appropriately pruned before pooling

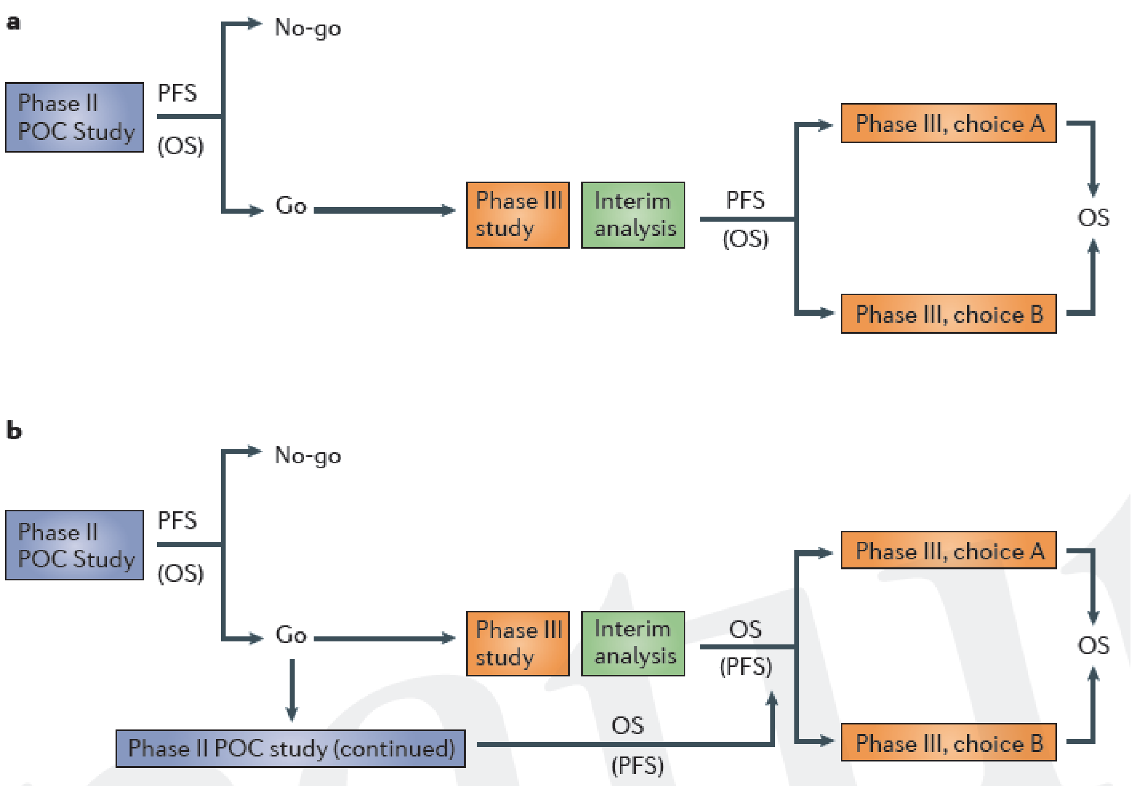
Challenge 2: Adjusting for Pruning

- Pruning indications that are doing poorly on surrogate endpoints may be seen as cherry picking
  - This can inflate the false positive rate, an effect termed "random high bias"
- Addressing the challenge:
  - Emphasize use of **external data**, such as maturing Phase 2 studies, for pruning
    - Pruning with external data does not incur a penalty for random high bias
  - Apply statistical penalty for control of type I error when applying pruning using **internal data**
    - Methods for calculating the penalty are described in stat methods papers (see key references)
    - Rules for applying penalty must be prospective
    - Penalty is not large enough to offset advantages of design

Challenge 3: Interim endpoints may not predict definitive endpoints

- Addressing the challenge:
  - Prefilter indications based on maturing definitve endpoint data from phase 2
    - See Figure 2
  - Require consistent trend in definitive endpoint for final full approval

Phase 2 Influencing Phase 3 Adaptation: The Phase 2+ Method



Beckman, R.A., Clark, J. & Chen, C. Integrating predictive biomarkers and classifiers into oncology clinical development programmes. *Nature Reviews Drug Discovery* **10,** 735-748 (2011)

Another Possible Source of External Data

- Real World Data (RWD) from Off-Label Use
- Impact of RWD on basket trial performance is currently under study in a project led by postdoctoral fellow Daphne Guinn

Challenge 4: Different Comparators

- Different arms may have different comparators
- Addressing the challenge:
  - Additive designs are most readily understood, ie A+X vs A, B+X vs B, etc.
  - For other designs, the central question is always the same: can X improve outcomes relative to SOC in a population positive for a biomarker indicating benefit from X

DETAILED DESIGN CONSIDERATIONS

Designs to Be Compared

- Sample size changes after pruning
  - D0: No pruning and no change (benchmark)
  - D1: No increase to sample size after pruning
  - D2: Sample size in pooled analysis after pruning remains same as planned for the trial (SS)
  - D3: Sample size for trial remains same after pruning as planned for the trial (SS)

Designs	Overall Trial	Pooled Population
D0	SS	SS
D1	<SS	<SS
D2	>SS	SS
D3	SS	<SS

Comparison of operating characteristics

- k=6 tumor indications with total planned event size (k*N*) ranging from 150-350
  - The true treatment effect is –log(0.6), or hazard ratio of 0.6 in a time-to-event trial
- Pruning occurs at when half of the events have occurred
- Number of active indications (g) with target effect size ranges from 3 to 6, with remaining ones inactive

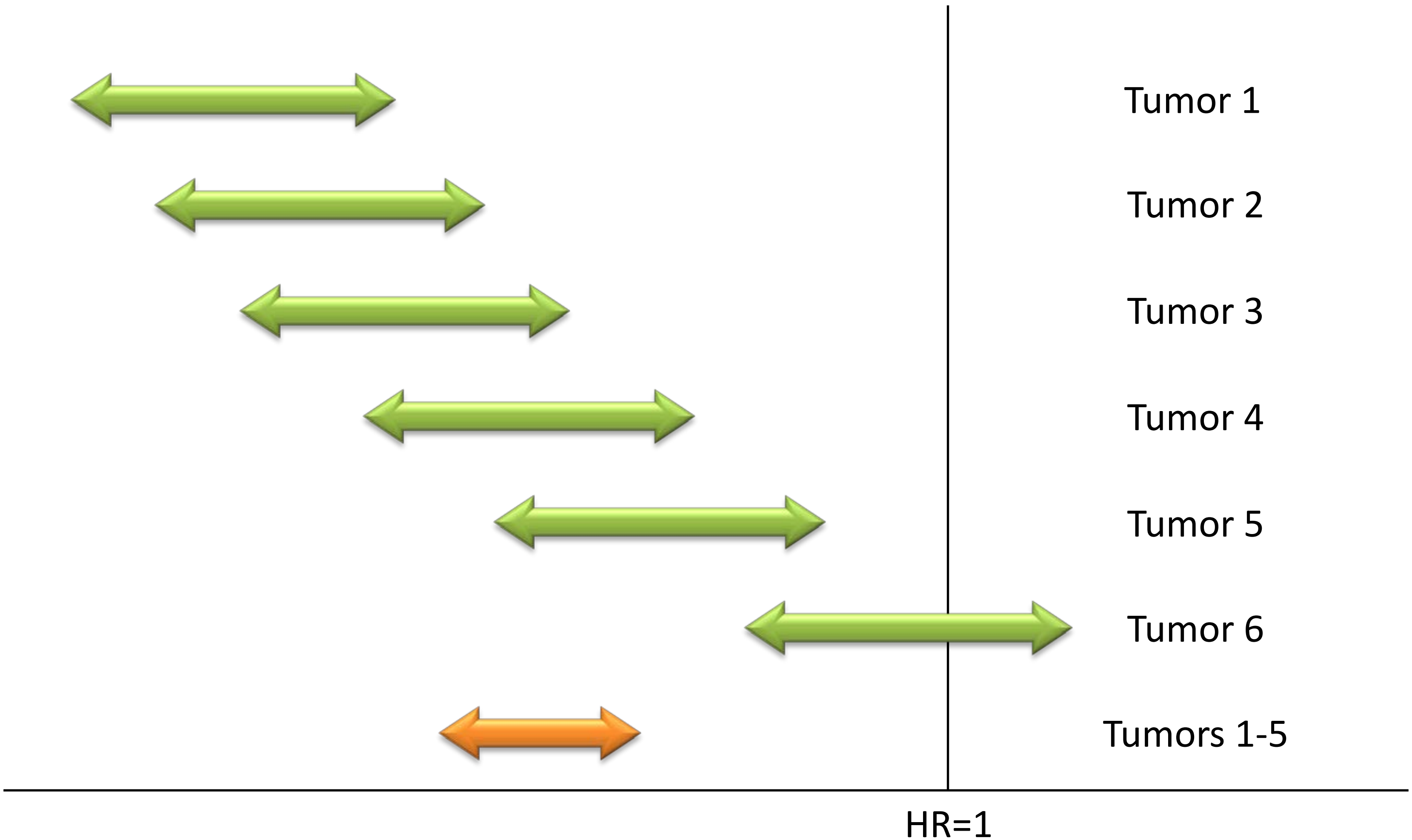
Study power and sample sizes under different pruning and pooling strategies

Planned events	Number of active tumors	Power (%) for a positive study				Exp. number of events for pooled population			Exp. number of events for overall study		
		D0	D1	**D2**	D3	D0/D2	D1	D3	D0/D3	D1	D2
200	6	95	85	**95**	93	200	157	179	200	179	221
200	5	85	75	**91**	86	200	144	172	200	172	228
200	4	67	62	**82**	76	200	131	166	200	166	234
200	3	44	45	**68**	61	200	119	159	200	159	240
300	6	99	96	**99**	99	300	254	277	300	277	323
300	5	96	81	**98**	96	300	232	266	300	266	334
300	4	84	81	**94**	91	300	209	255	300	255	345
300	3	60	64	**84**	79	300	187	244	300	244	356

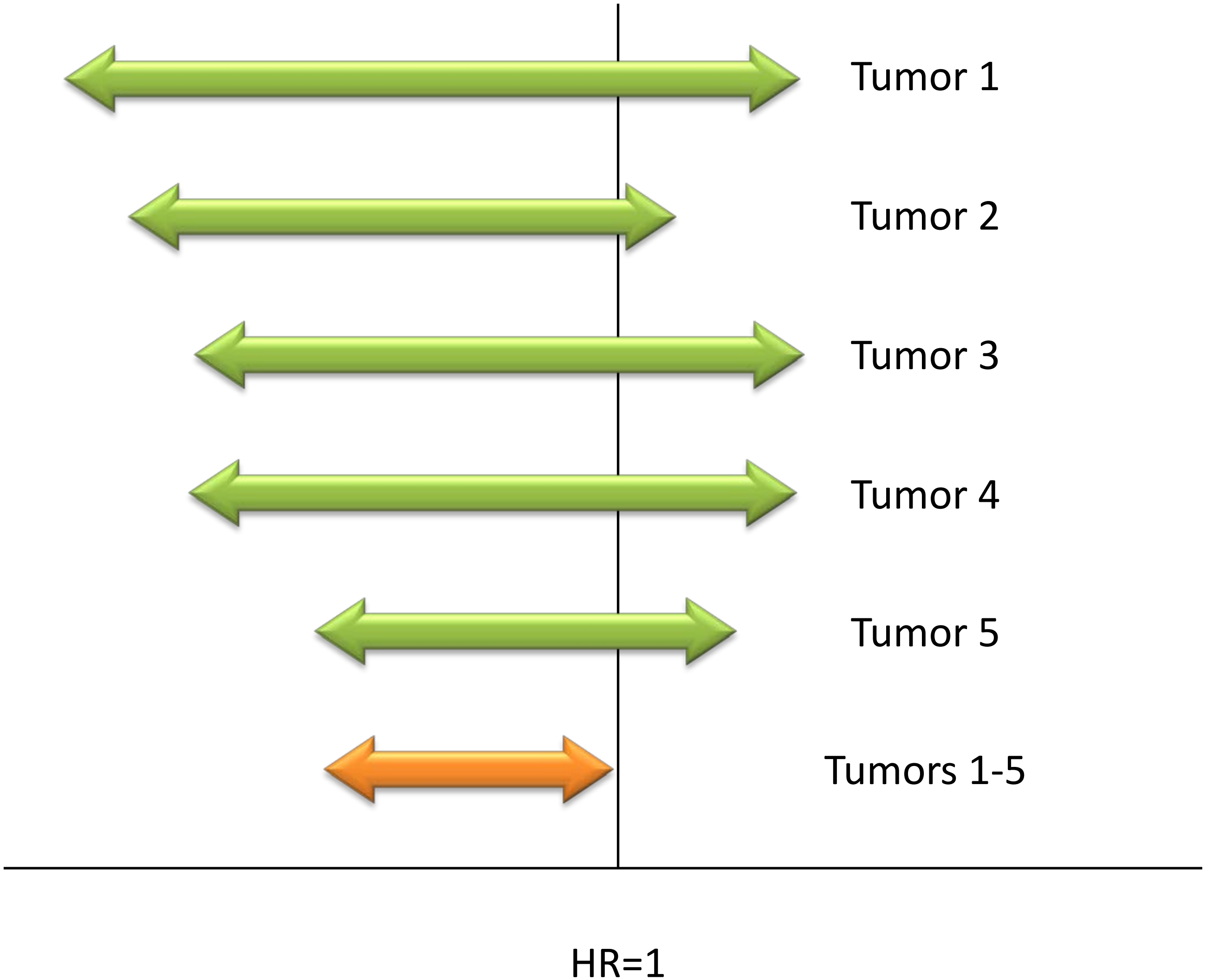
An Application of Special Interest

- A randomized controlled basket trial with 1:1 randomization in 6 tumor indications, each targeting a hazard ratio of 0.5 in PFS with 90% power at 2.5% alpha
  - 88 PFS events and 110 patients planned for each indication
  - PFS analysis is conducted when all are enrolled
- D2 is applied to keep total sample size at 660 in pooled population targeting 430 death events
  - The study has ~90% power to detect a hazard ratio of 0.7 in OS at 0.8% alpha (after taking the penalty) assuming ρ=0.5
  - Observed hazard ratio ~0.79 or lower for a positive trial in pooled population (vs ~0.84 under D0)
- Potential to gain approvals in 6 indications based on comparable sample size to a conventional Phase 3 trial

Forest Plot of Hypothetical PFS Outcome



Forest Plot of Hypothetical OS Outcome



Compare to a Conventional Confirmatory Trial

Conventional

Inconsistency in treatment effect (e.g., gender, age) is not suspected

↓

Primary analysis in pooled population, and, if positive, consistency will be assessed ad hoc

Basket (two-stage)

Consistency in treatment effect across tumor indication is less certain

↓

Inactive ones are pruned at an interim analysis

↓

Primary analysis in pooled population of remaining ones, and, if positive, consistency will be assessed ad hoc

Limitations

- The design as presented is not applicable to ultra rare indications or indications with no standard of care
  - Randomization is not feasible/ethical under these conditions
- Type I error control is currently limited to the global null hypothesis

Ongoing Research

- Control of family-wise Type I error rate
  - It will likely be feasible to control it to the Type I error which would be inherent in separate trials
- Incorporation of Real World Data (RWD) for indication and endpoint selection
  - Real world data not used for approval in our research

Conclusions

- It is feasible to create a general design concept for a basket study that is suitable for effective agents
- Multiple challenges can be addressed with careful planning
- Benefits include:
  - Increased and earlier patient access to targeted therapies for small subgroups
  - Cost-effective methods for sponsors to develop targeted agents in small subgroups
  - More robust datasets for health authorities to assess benefit-risk in these small patient groups

Key References

- Li, Wen, Chen, Cong, Li, Xiaojun, and Beckman, Robert A. Estimation of treatment effect in two-stage confirmatory oncology trials of personalized medicines. Statistics in Medicine, in press (2017).
- Beckman, Robert A., Antonijevic, Zoran, Kalamegham, Rasika, and Cong Chen. Adaptive Design for a Confirmatory Basket Trial in Multiple Tumor Types Based on a Putative Predictive Biomarker. Clinical Pharmacology and Therapeutics, 100: 617-625 (2016).
- Yuan, Shuai S, Chen, Aiying, He, Li, Chen, Cong, Gause, Christine K, and Robert A. Beckman. On Group Sequential Enrichment Design for Basket Trial. Statistics in Biopharmaceutical Research, 8: 293-306 (2016).
- Chen, Cong, Li, Nicole, Yuan, Shuai, Antonijevic, Zoran, Kalamegham, Rasika, and Robert A. Beckman. Statistical Design and Considerations of a Phase 3 Basket Trial for Simultaneous Investigation of Multiple Tumor Types in One Study. Statistics in Biopharmaceutical Research, 8: 248-257 (2016).
- Magnusson BP, Turnbull BW. Group sequential enrichment design incorporating subgroup selection. *Stat Med*. 2013;32(16):2695-2714.
- Heinrich MC, Joensuu H, Demetri GD, Corless CL, Apperley J, Fletcher JA, et al. Phase II, open-label study evaluating the activity of imatinib in treating life-threatening malignancies known to be associated with imatinib-sensitive tyrosine kinases. *Clin Cancer Res*. 2008;14(9):2717-2725.
- Demetri G, Becker R, Woodcock J, Doroshow J, Nisen P, Sommer J. Alternative trial designs based on tumor genetics/pathway characteristics instead of histology. Issue Brief: Conference on Clinical Cancer Research 2011; http://www.focr.org/conference-clinical-cancer-research-2011.

Backup Slides

Challenge 3: Will the companion diagnostic assay generalize across indications?

- Analytical properties of assay may depend on tissue type
- Cutoff between biomarker positive and negative may vary between tissue types for a continuous biomarker
- Addressing the challenge:
  - Analytical validation of the assay for all relevant indications prior to study start
  - Prior to study start, recommend biomarker stratified randomized phase 2 studies to set provisional cutoffs for continuous biomarkers in each indication to the extent feasible

Challenge 4: Availability of tissue

- Tissue sampling and processing are variables that can greatly affect the outcome of a study based on a predictive biomarker
- Basket studies will require cooperation and uniformity across departments organized by histology
- Addressing the challenge:
  - The sponsor must have extensive contact with the pathology department and relevant clinical departments at all investigative sites and provide standard methods for tissue sampling, handling, and processing
  - The sponsor should engage an expert pathologist who is dedicated to training prior to trial start, and troubleshooting during the trial

Type I error control

- k tumor indications each with sample size of N and all with 1:1 randomization
- An interim analysis is conducted at information fraction t for each tumor indication and a tumor will not be included in the pooled analysis if p-value>α_t
- The pooled analysis will be conducted at α^* so that the overall Type I error is controlled at α when there is no treatment effect for any tumor (H0)
- What is α^*?

Solving for adjusted alpha (α^*)

- Let Y_{i1} be the test statistics based on information fraction *t*, and Y_{i2} be the test statistics based on the final analysis of data in the *i*-th cohort (i=1, 2,…,k)
- Suppose that *m* cohorts are included in the final analysis ($m \geq 1$), and let V_m be the corresponding test statistics. The probability of a positive outcome in pooled analysis is

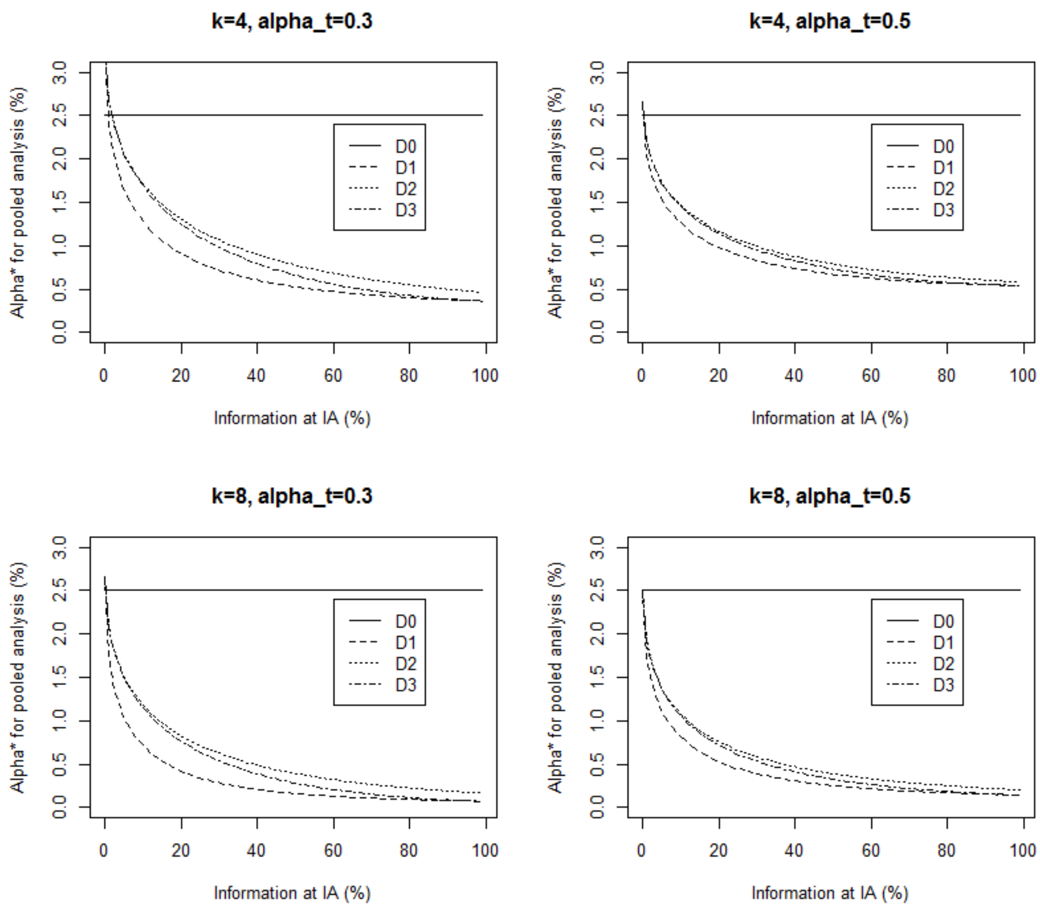
$$Q_0(\alpha^*|\alpha_t, m) = \Pr_{H_0}(\cap\{Y_{i1} > Z_{1-\alpha_t} \text{ for } i=1,\ldots,m\}, \cap\{Y_{j1} < Z_{1-\alpha_t} \text{ for } j=m+1,\ldots k\}, V_m > Z_{1-\alpha^*})$$

or $$Q_0(\alpha^*|\alpha_t, m) = \Pr_{H_0}(\cap\{Y_{i1} > Z_{1-\alpha_t} \text{ for } i=1,\ldots,m\}, V_m > Z_{1-\alpha^*})(1-\alpha_t)^{(k-m)}$$

- α^* is solved from below where $c(k, m) = k!/((k-m)!m!)$

$$\sum_{m=1}^{k} c(k,m) Q_0(\alpha^*|\, \alpha_t, m) = \alpha$$

α* under different design options



α* decreases with increasing k as expected, but its relationship with α_t is complicated with the interplay between cherry-picking and futility stopping.