A Process View of Statistics for Clinical Trials: ANOVA, Product-Limit, and Adaptive Design

Xiaolong Luo, Ph.D., MBA
Mingyu Li, Ph.D. and S. Peter Ouyang, Ph.D.

Celgene Corporation

Seventeenth Annual Biopharmaceutical Applied Statistics Symposium
November 8 - 12 2010
Hilton Oceanfront Resort
Hilton Head Island, South Carolina

Outline

- Inherited pros and cons from Fisherian statistics for clinical trials
- Extension of backend dynamic for subject departure: survival analysis
- Deficiency of some adaptive designs
- BMPP framework for clinical trials
- Applications to adaptive designs
- Statistical view of operational bias
- Conclusions

Fisher's Prototype Example

- Outcome: plant yield in pounds
- Treatments:
  - Basal only
  - Sulphate
  - Chloride
- Questions
  - Manuring effect
  - Variety effect
  - Plot effect
- Statistical Method
  - ANOVA

TABLE 46

	Sulphate Row.	Sulphate Row.	Sulphate Row.	Chloride Row.	Chloride Row.	Chloride Row.	Basal Row.	Basal Row.	Basal Row.
Ajax	3·20	4·00	3·86	2·55	3·04	4·13	2·82	1·75	4·71
Arran Comrade	2·25	2·56	2·58	1·96	2·15	2·10	2·42	2·17	2·17
British Queen	3·21	2·82	3·82	2·71	2·68	4·17	2·75	2·75	3·32
Duke of York	1·11	1·25	2·25	1·57	2·00	1·75	1·61	2·00	2·46
Epicure	2·36	1·64	2·29	2·11	1·93	2·64	1·43	2·25	2·79
Great Scot	3·38	3·07	3·89	2·79	3·54	4·14	3·07	3·25	3·50
Iron Duke	3·43	3·00	3·96	3·33	3·08	3·32	3·50	2·32	3·29
K. of K.	3·71	4·07	4·21	3·39	4·63	4·21	2·89	4·20	4·32
Kerr's Pink	3·04	3·57	3·82	2·96	3·18	4·32	2·00	3·00	3·88
Nithsdale	2·57	2·21	3·58	2·04	2·93	3·71	1·96	2·86	3·56
Tinwald Perfection	3·46	3·11	2·50	2·83	2·96	3·21	2·55	3·39	3·36
Up-to-Date	4·29	2·93	4·25	3·39	3·68	4·07	4·21	3·64	4·11

The First RCT

BRITISH MEDICAL JOURNAL

LONDON SATURDAY OCTOBER 30 1948

STREPTOMYCIN TREATMENT OF PULMONARY TUBERCULOSIS

A MEDICAL RESEARCH COUNCIL INVESTIGATION

The following gives the short-term results of a controlled investigation into the effects of streptomycin on one type of pulmonary tuberculosis. The inquiry was planned and directed by the Streptomycin in Tuberculosis Trials Committee, composed of the following members: Dr. Geoffrey Marshall (chairman), Professor J. W. S. Blacklock, Professor C. Cameron, Professor N. B. Capon, Dr. R. Cruickshank, Professor J. H. Gaddum, Dr. F. R. G. Heaf, Professor A. Bradford Hill, Dr. L. E. Houghton, Dr. J. Clifford Hoyle, Professor H. Raistrick, Dr. J. G. Scadding, Professor W. H. Tytler, Professor G. S. Wilson, and Dr. P. D'Arcy Hart (secretary). The centres at which the work was carried out and the specialists in charge of patients and pathological work were as follows:

The Control Scheme

Determination of whether a patient would be treated by streptomycin and bed-rest (S case) or by bed-rest alone (C case) was made by reference to a statistical series based on random sampling numbers drawn up for each sex at each centre by Professor Bradford Hill; the details of the

TABLE II.—*Assessment of Radiological Appearance at Six Months as Compared with Appearance on Admission*

Radiological Assessment	Streptomycin Group		Control Group	
Considerable improvement ..	28	51%	4	8%
Moderate or slight improvement	10	18%	13	25%
No material change	2	4%	3	6%
Moderate or slight deterioration	5	9%	12	23%
Considerable deterioration ..	6	11%	6	11%
Deaths	4	7%	14	27%
Total	55	100%	52	100%

Statistical Science for Clinical Trials

- Estimation of treatment effect to decide risk and benefit ratio of any new treatment
  - Ranking of treatment effects is not sufficient
  - Fisher (1922) on complete theoretical treatment of data: parameter for specifying population; its estimation from sample; exact form of distribution of the statistics
- Test of significance
  - Fisher (1925) Statistical Methods for Research Workers
- Randomization
  - Validate distribution theory

Difference in the Conduct of Experiments

	Clinical Trial	Potato Yield
Subject arrival	Sequential	Batch
Subject departure	Variable with possible missing outcome measure	Constant with definite outcome measure
Ethical concern	• Patients in general • Patients as experimental units	None
Length of experiment	Potentially long and unpredictable	Relatively short and predictable
Change of intervention	Often unavoidable	Unlikely
Mid course action	Highly desirable for ethical concerns	May not be an issue

Clash between Clinical Trials and Fisher's Data Modeling

- **Outcome distribution may not be solely determined by treatment due to**
  - **Dynamic accrual carrying information other than treatment alone**
  - **Dynamic departure leading to partially observable data**
  - **Ethical concerns altering experiment course**
- **Estimation and significance test may not be justified due to tainted intervention and missing outcome data**

Survival Analysis: Backend Dynamic

- Kaplan and Meier (1958)

$$\hat{P}(t) = \prod_{r} \frac{N - r}{N - r + 1}$$

- Cox (1972)

$$l(\beta) = \prod_{i} \frac{e^{\beta X_{(i)}}}{\sum_{Y_j \geq Y_{(i)}} e^{\beta X_j}}$$

- Aalen (1978)

$$M_i(t) = 1(Y_i \leq t, \delta_i = 1) - \int_0^t 1(Y_i \geq s) h_i(s) ds$$

Dynamic Subject Accrual

- Armitage (1960) Sequential medical Trials
- Box and Jenkins (1962) Some Statistical Aspects of Adaptive Optimization
  - Empirical feedback vs.technical feedback
- Zelen (1969) Play the Winner Rule and the Controlled Clinical Trial
- Pocock and Simon (1975) Sequential Treatment Assignment with Balancing Prognostic Factors in the Controlled Clinical Trials

Recent Development of Adaptive Design

- L. Fisher (1998) Self-designing clinical trials
- Cui, Hung, Wang (1999) Modification of sample size in group sequential clinical trials
- Liu and Chi (2001) On sample size and inference for two-stage adaptive design
- Muller and Schafer (2001) Adaptive group sequential designs for clinical trials
- EMEA 2006: Reflection paper on methodological issues in confirmatory clinical trials with flexible design and analysis plan
- FDA 2010: Adaptive design clinical trials for drugs and biologics
- …

Methodological Deficiency of Some Adaptive Designs

1. Lack of proper estimation for treatment effect
   - Testing alone may fall into Fisher's concerns on Neyman-Pearson theory
   - Treatment selection based on relative efficacy ranking is not sufficient
2. May require instantaneous observations
   - Most clinical trial outcome measures take time to assess
3. Statistics may not adjust for complicated adaptive rules (e.g., PWR and dynamic randomization) to provide a convincing test of the null hypothesis
4. May assume normal distribution or known variance for the outcome variable

"It is clear that statistical literature on adaptive treatment assignment has had little impact on the conduct of clinical trials."

Simon (1977) Biometrics 33, 743-749

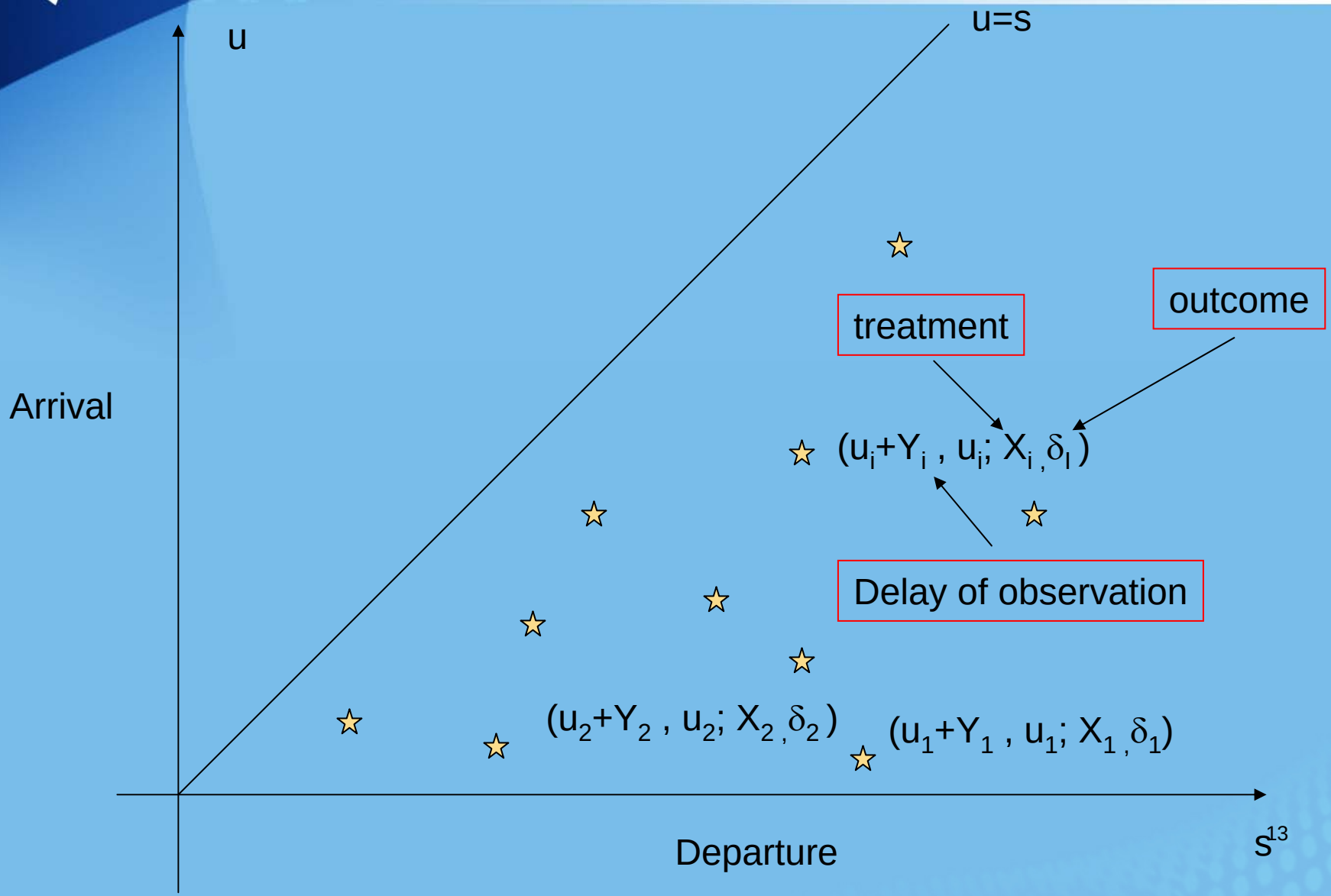
How can we solve those problems ?

子曰: 工欲善其事,必先利其器

Sharpening the tool
before doing the work !!!

By Confucius

A BMPP Framework for Clinical Trial



Random Field and Bivariate Marked Point Process

- A random sequence $(T_n, X_n)_{n\geq 1}$ is a (univariate) marked point process (Last and Brandt 1995)
- Cox and Lewis (1975) introduced bivariate point process recording two types of events such as subject's arrival and departure
- Jacod and Shiryaev (1987, 2002) showed existence of compensator for integer valued random measure, yielding martingale measure
- Kallianpur and Xiong (1995) described stochastic integral with respect to martingale measure

Bivariate Marked Point Process

$$R_t = \sum_i 1(u_i \le t)$$

$$p_a(dudx) = 1(X_u \in dx)dR_u$$

$$p(dsdudxd\delta) = 1(X_u \in dx, Y_u \in ds)dR_u d\delta$$

$$\Im_t = \sigma(p_a(dudx), p(dsdudxd\delta) : u \le s \le t, x, \delta)$$

$$q(dsdudxd\delta) = E[1(X_u \in dx, Y_u \in ds)dR_u d\delta \mid \Im_{s-}]$$

$$dM_s = p(dsdudxd\delta) - q(dsdudxd\delta)$$

BMPP and Adaptive Design

- **Adapt the enrollment process R_t for treatment and/or population selection**

$$r_s(x) = \begin{cases} \frac{1}{K+1} & \forall\, x,\ 0 < s < t_1; \\ \frac{1}{K+1} & \forall\, x,\ t_1 \le s < T,\ \omega \notin C_1; \\ \frac{1}{2} & x = 0, x_1,\ t_1 \le s < T,\ \omega \in C_1; \\ 0 & \text{o.w.} \end{cases}$$

- **Use the predictable stopping time to adapt sample size**

$$\tau_2 = \inf\{t > \tau_1 : \int_{[\tau_1, t]\times\{0,1,\ldots,K\}\times R} g(s,x,\delta;\omega) q(dsdxd\delta) > I(\tau_1)\}$$

Statistical Inference with BMPP

$$\frac{\int_{[0,t]\times\{0,1,\ldots,K\}\times R} f(s,x,\delta;\omega)\,p(dsdxd\delta) - \int_{[0,t]\times\{0,1,\ldots,K\}\times R} f(s,x,\delta;\omega)q(dsdxd\delta)}{\sqrt{E\int_{[0,t]\times\{0,1,\ldots,K\}\times R} f(s,x,\delta;\omega)^2\,q(dsdxd\delta)}} \sim n(0,1)$$

- **The distribution is good for any predictable enrollment process and treatment assignment, validating the significance test**
- **Can be used as an estimating equation for parameter estimation**

How Does BMPP Correct Those Methodological Deficiency from Some Adaptive Designs?

1. Provide estimating equation for treatment effect
2. Naturally accommodate delayed observations
3. Validate probability distribution of statistics adjusted for any adaptive rules
4. No need for parametric modeling under regularity conditions

VOLUME 76 • OCTOBER 1985 • NUMBER 4

Pediatrics

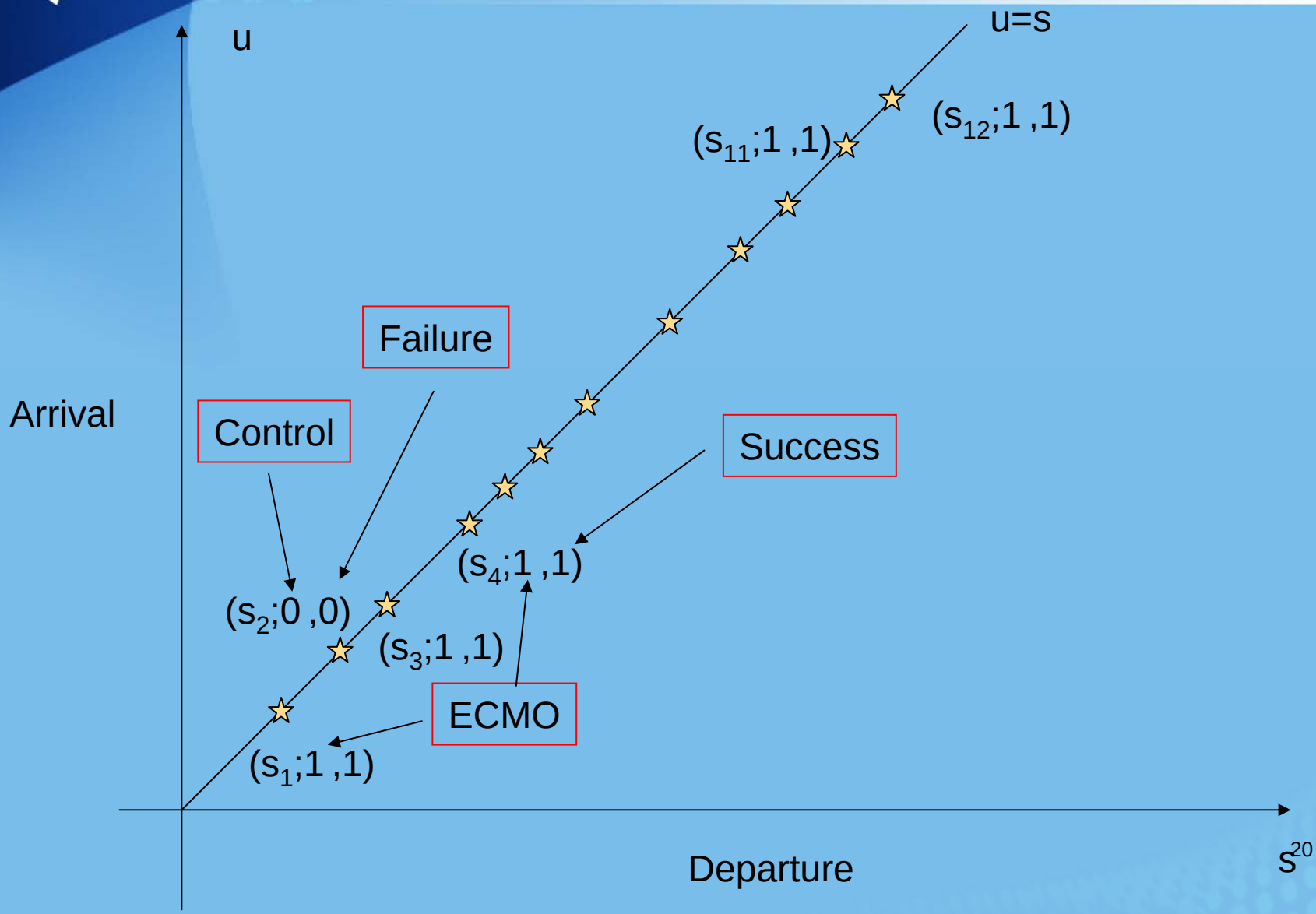
Extracorporeal Circulation in Neonatal Respiratory Failure: A Prospective Randomized Study

Robert H. Bartlett, MD, Dietrich W. Roloff, MD, Richard G. Cornell, PhD, Alice French Andrews, MD, Peter W. Dillon, MD, and Joseph B. Zwischenberger, MD

The first patient was randomly assigned to ECMO and survived. The second patient was randomly assigned to conventional treatment and died. Hence, the odds of the next patient being randomly assigned to ECMO were 3:1. The next patient was randomly assigned to ECMO and survived. This pattern continued until there were ten who had been treated with ECMO or ten control patients who died. This pattern was established in accord

domly assigned to that treatment. A group of 12 infants with birth weight greater than 2 kg met objective criteria for high mortality risk. One patient was randomly assigned to conventional treatment (that patient died); 11 patients were randomly chosen for extracorporeal membrane oxygenation (all survived). Intracerebral hemorrhage occurred in one of 11 surviving children. Extracorporeal membrane oxygenation allows lung rest and improves survival compared to conventional ventilator therapy in newborn infants with severe respiratory failure. *Pediatrics* 1985;76:479–487; *neonatal, respiratory failure, extracorporal circulation, oxygenation.*

ECMO Trial


u
u=s
$(s_{12};1,1)$
$(s_{11};1,1)$
Failure
Arrival
Control
Success
$(s_4;1,1)$
$(s_2;0,0)$
$(s_3;1,1)$
ECMO
$(s_1;1,1)$
Departure
s^{20}

Marked Point Process for ECMO Trial

$$R_t = \sum_i 1(s_i \le t)$$

$$\pi_x = P(\delta_i = 1 \mid X_i = x), x = 0,1$$

$$p(dsdxd\delta) = 1(X_s \in dx, \delta_s = \delta)dR_s$$

$$n(t, j) = \int_0^t 1(X_s = j)dR_s$$

$$B(t) = \int_0^t 1(\delta_s = 1)dR_s, B(t, j) = \int_0^t 1(X_s = j, \delta_s = 1)dR_s$$

$$r_s = P(X_s = 1 \mid F_{s-}) = \frac{2B(s-,1) + 1 + R_{s-} - n(s-,1) - B(s-)}{R_{s-} + 2}$$

$$q(dsdxd\delta) = [xr_s + (1-x)(1-r_s)]\pi_x{}^{\delta}(1-\pi_x)^{1-\delta} dR_s dxd\delta$$

$$dM_s = p(dsdudxd\delta) - q(dsdudxd\delta)$$

$$\tau = \inf\{t : \max\{n(t,1), n(t,0)\} \ge 10\}$$

Asymptotic Pivot

$$\frac{\iint_{[0,t]\times X} f(s,x,\delta;\omega)\, p(dsdxd\delta) - (\pi_1 - \pi_0)\int_0^t r_s(1-r_s)dR_s}{\sqrt{\iint_{[0,t]\times X} f(s,x,\delta;\omega)^2\, p(dsdxd\delta)}} \sim n(0,1)$$

$$f(s,x,\delta;\omega) = [1(x=1) - r_s]1(\delta = 1)$$

$$\iint_{[0,t]\times X} f(s,x,\delta;\omega) q(dsdxd\delta) = (\pi_1 - \pi_0)\int_0^t r_s(1-r_s)dR_s$$

$$E\iint_{[0,t]\times X} f(s,x,\delta;\omega)^2 q(dsdxd\delta)$$

$$= \pi_0 E\int_0^t 1(\delta_s = 1) r_s{}^2(1-r_s)dR_s + \pi_1 E\int_0^t 1(\delta_s = 1) r_s(1-r_s)^2 dR_s$$

Test of Significance

$$\frac{\iint_{[0,\tau]\times X} f(s,x,\delta;\omega)\, p(dsdxd\delta)}{\sqrt{\iint_{[0,\tau]\times X} f(s,x,\delta;\omega)^2\, p(dsdxd\delta)}} \sim n(0,1)$$

Table 1: ECMO Trial Data and MPP Statistics

s	s_1	s_2	s_3	s_4	...	s_{11}
R_{s-}	0	1	2	3	...	10
X_s	1	0	1	1	...	1
δ_{s-}	1	0	1	1	...	1
$n(s-,1)$	0	1	1	2	...	9
$B(s-,1)$	0	1	1	2	...	9
$B(s-)$	0	1	1	2	...	9
r_s	$\frac{1}{2}$	$\frac{2}{3}$	$\frac{3}{4}$	$\frac{4}{5}$	...	$\frac{11}{12}$
$f(s,)$	$\frac{1}{2}$	0	$\frac{1}{4}$	$\frac{1}{5}$	...	$\frac{1}{12}$

$$T_{ECMO} = \frac{\frac{1}{2} + \sum_{i=4}^{12} \frac{1}{i}}{\sqrt{\frac{1}{2^2} + \sum_{i=4}^{12} \frac{1}{i^2}}} = 2.627$$

P = 0.0086

Estimation of Treatment Effect

$$\hat{\Delta} = \frac{\iint_{[0,\tau]\times X} f(s,x,\delta;\omega)\, p(dsdxd\delta)}{\int_0^{\tau} r_s(1-r_s)dR_s}$$

$$Var(\hat{\Delta}) = \frac{\iint_{[0,\tau]\times X} f(s,x,\delta;\omega)^2\, p(dsdxd\delta)}{\left[\int_0^{\tau} r_s(1-r_s)dR_s\right]^2}$$

$$ECMO:$$

$$\hat{\Delta} = \frac{\frac{1}{2} + \sum_{i=4}^{12} \frac{1}{i}}{\frac{1}{2^2} + \sum_{i=4}^{12} \frac{1}{i}\frac{i-1}{i}} = 1.34$$

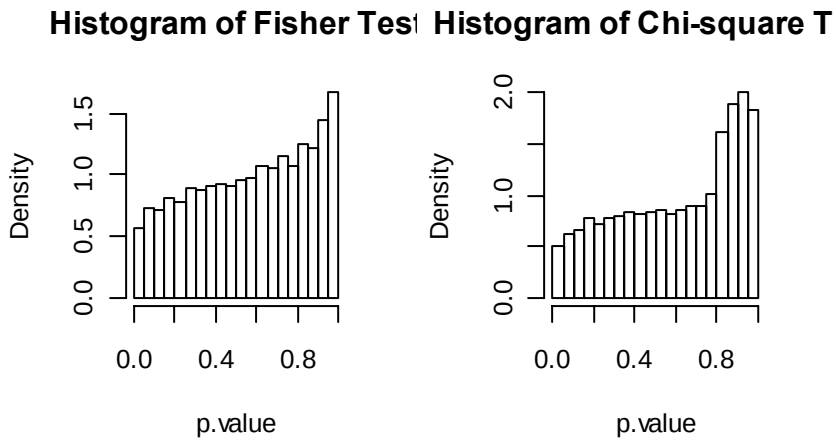
$$Var(\hat{\Delta}) = 0.26$$

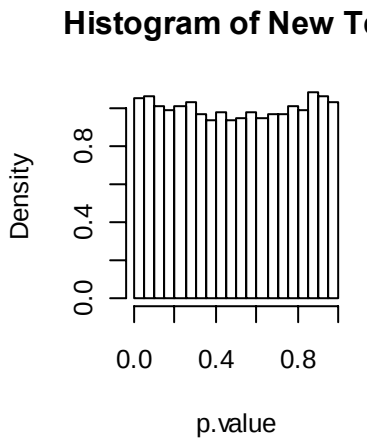
$$sd = 0.51$$

Cautious with small sample size !

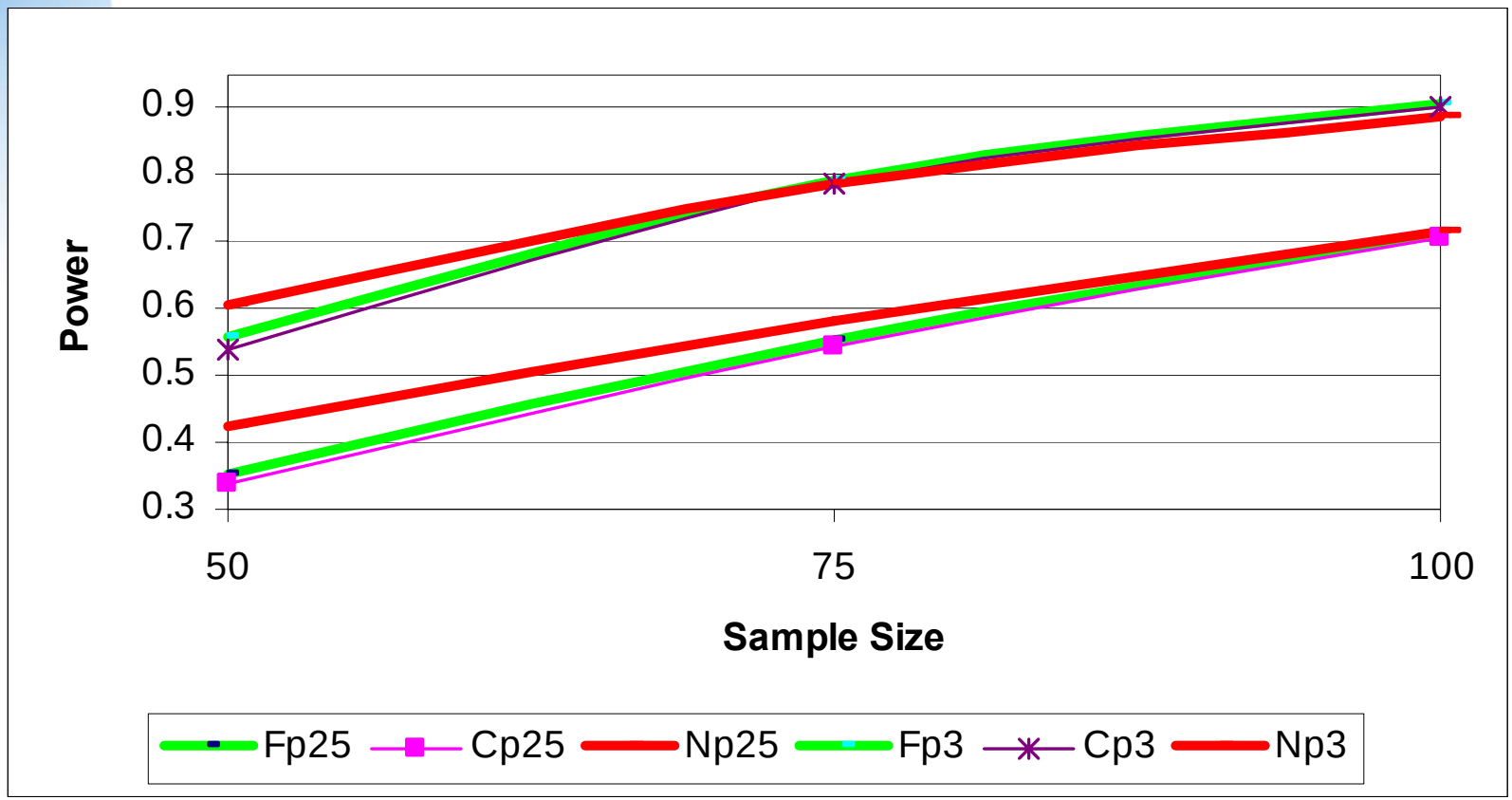
Simulations Under Null

- n = 50, N = 20,000
- $P_0 = P_1 = 0.1$
- Proportion with p < 0.025
  - 0.009 for the Fisher naïve
  - 0.015 for Chi square naïve
  - 0.026 for the new





Simulations Under Alternatives



Based on N = 20, 000. 0.1 vs 0.25 and 0.1 vs 0.3.

Some Potential Bias from Trial Operations

- We assume that the outcome distribution depends on only the given treatment
  - Can investigators systematically change the distribution without changing the treatment?
  - What may lead investigators to manage patients differently in a systematic manner?
- We assume that the treatment assignment depends on only the randomization system at the given time
  - What may prompt investigators to select subjects for randomization?
- We assume the distribution of subjects available for the trial accrual does not change over time
  - Can investigators encourage or discourage some type of subjects differently over time?
- We assume that the trial cutoff is a predictable stopping time
  - Can it be decided without satisfying the predictability?

Bias Minimization Plan

- Keep the distribution of the interim study results on a need-to-know basis
- Be cautious on investigators being influenced outside the protocol; including even external information that may alter the basic assumptions
- Predictable adaptation does not have to bias the study as long as the operations are as planned, e.g., using IVRS to blind the process
- Build firewall between unblinded DMC and project team by executive committee
- Ensure adequate sample size for normal approximation

Conclusions

- Fisher's static modeling framework may not be appropriate for applications in clinical trials with dynamic data flow, however the concepts of estimation and test of significance are still critically important
- Survival analysis is one important extension to allow dynamic subject departure in the traditional static model
- The proposed BMPP framework removes the static modeling limitation but still answers the same statistical questions for clinical trials
  - Naturally addressing problems associated with adaptive design
- Implementation for Play the Winner Rule (PWR) design is illustrated via ECMO trial
  - The BMPP procedure is more robust in moderate sample size and comparable with naïve procedures in large sample sizes
- Further application to adaptive treatment selection, population enrichment, sample size re-estimation, dynamic randomization and other adaptive designs can be done similarly and will be presented separately