Developing Implementable Bandit-Based Designs for Clinical Trials: Where Methods Meet Practice

Sofía S. Villar

Workshop on Multi-Armed Bandits and Learning Algorithms
Erasmus University Rotterdam, 25 May 2018.

Outline

Optimality in Clinical Trials

Theory and Practice

Gittins & Jones, 1979 Biometrika:

"*T*he two-armed bandit problem is so-called because it models the situation faced by a gambler using a fruit machine with two arms, instead of just one. When an arm is pulled the result is that the gambler either wins a prize or not. [...]The gambler's problem is to choose a sequence of pulls on the two arms, which depends in a sequential manner on the record of successes and failures, in such a fashion as to maximize his expected total gains. [...] Multi-armed bandit problems (MABP) are similar, but with more than two arms. **Their chief practical motivation comes from clinical trials**, though they are also of interest as probably the simplest worthwhile set of problems in the sequential design of experiments."

Almost 40 years later, these ideas remain largely unused in practice.

Optimality in Clinical Trials

Theory and Practice

Gittins & Jones, 1979 Biometrika:

"*T*he two-armed bandit problem is so-called because it models the situation faced by a gambler using a fruit machine with two arms, instead of just one. When an arm is pulled the result is that the gambler either wins a prize or not. [...]The gambler's problem is to choose a sequence of pulls on the two arms, which depends in a sequential manner on the record of successes and failures, in such a fashion as to maximize his expected total gains. [...] Multi-armed bandit problems (MABP) are similar, but with more than two arms. **Their chief practical motivation comes from clinical trials**, though they are also of interest as probably the simplest worthwhile set of problems in the sequential design of experiments."

Almost 40 years later, these ideas remain largely unused in practice.

The Learning-Earning Dilemma in Clinical Trials

An ethical Problem

Even if their scope is much more general, the most common scenario chosen to motivate the MABP across seminal papers is that of a *clinical trial* assumed to aim at balancing two separate goals:

G1 To correctly identify the best treatment (learning).

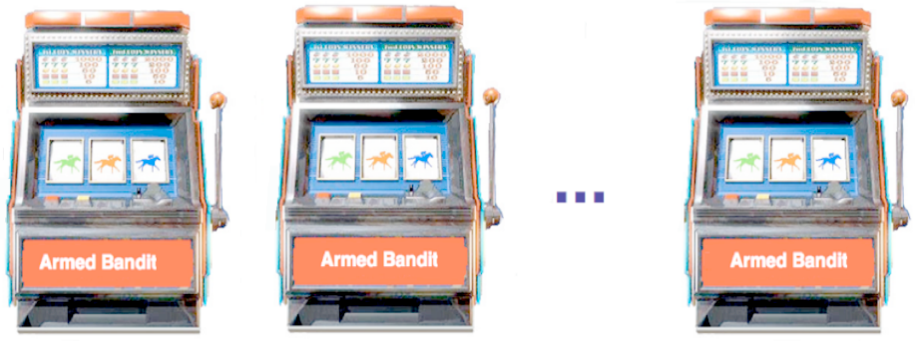
G2 *To most effectively treat as many patients as possible (earning).*

The ethical conflict around these goals is always present but it becomes more acute (*suboptimality gap grows*) when: the population with a disease is small, the disease is life-threatening and/or there are several potential treatments to study at once.

Traditional trial design is focused on controlling error probabilities (*learning*) and it largely ignores (*earning*): patient horizon and optimising treatment for patients in the trial (and for the population as a whole).

Outline

The (*Classic*) Multi-armed Bandit Problem

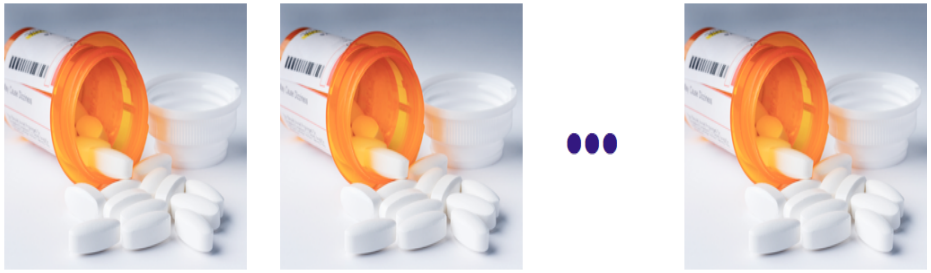


$£Y_{1,t}$ $£Y_{2,t}$ $\ldots$ $£Y_{k,t}$

Maximise total expected gains over time:

learn about the success rates of the slot machines
just enough to maximise average total profit

Trial design as a (*classic*) Multi-armed Bandit Problem



$Y_{1,t}$ $Y_{2,t}$ $\ldots$ $Y_{k,t}$

Maximise total expected *patient benefit* over time:

learn about the treatments' efficacy just enough to maximise patients' outcomes over the population

The Curse of Dimensionality

MABP and Computational Feasibility

- Solution to the MABP according to Bellman's principle of optimality exists but is computationally expensive. Prohibitively so for most realistic scenarios.

- The curse of dimensionality was (till early 80's) the single most important limitation to its applicability in practice (in any context).

Armitage (1985):

"*The problem can now be seen as essentially the 'two-armed bandit' problem for a finite horizon. The solution to this can in principle be obtained by dynamic programming methods, but in practice the computation involved is prohibitive except for trivially small horizons.*"

The Gittins Index for a Clinical Trial

Beyond the Computational Limitation...

Gittins (1979)
"*Gittins index rule: divide and conquer strategy*"

Despite being computationally feasible for multi-armed trials (and simpler than DP to summarise), index rules have not been applied to a trial yet.

Important barriers to its use in practice include (Villar et al, 2015a):

(1) Its fully sequential nature: outcomes must be immediately available.

(2) Decisions are not randomized: treatment allocation bias, covariate imbalance. Basis for inference

(3) Given an objective degree of discrimination between two treatments, it lacks a sufficient/comparable level of statistical power.

(4) It does not incorporate potentially important prognostic covariates.

(5) Others: bias in estimation of treatment effect (overestimation of treatment effect), the effect of *patient drift*, etc.

The Gittins Index for a Clinical Trial

Beyond the Computational Limitation...

Gittins (1979)
"*Gittins index rule: divide and conquer strategy*"

Despite being computationally feasible for multi-armed trials (and simpler than DP to summarise), index rules have not been applied to a trial yet.

Important barriers to its use in practice include (Villar et al, 2015a):

(1) Its fully sequential nature: outcomes must be immediately available.

(2) Decisions are not randomized: treatment allocation bias, covariate imbalance. Basis for inference

(3) Given an objective degree of discrimination between two treatments, it lacks a sufficient/comparable level of statistical power.

(4) It does not incorporate potentially important prognostic covariates.

(5) Others: bias in estimation of treatment effect (overestimation of treatment effect), the effect of *patient drift*, etc.

Outline

The Forward Looking Gittins Index

Introducing Randomization to the Gittins Index Rule

Assume that T patients are arrive sequentially in blocks of size b over J stages, so that $J \times b = T$. In Villar et al (2015b) we defined group allocation probabilities based on the Gittins Index (GI) rule as follows:

Simplest example: $b = 2$. Priors: control $(s_{(0,0)}, f_{(0,0)}) = (1, 2)$ and experimental $(s_{(1,0)}, f_{(1,0)}) = (1, 1)$

$$j = 1,$$
$$\mathcal{G}_1(1,1) = 0.8699$$
$$\mathcal{G}_0(1,2) = 0.7005$$

$\frac{1}{2}$ / $Y_{1,0} = 1$

$$j = 2,$$
$$\mathcal{G}_1(2,1) = 0.9102$$
$$\mathcal{G}_0(1,2) = 0.7005$$

$\frac{1}{2}$ / $Y_{1,0} = 0$

$$j = 2,$$
$$\mathcal{G}_1(1,2) = 0.7005$$
$$\mathcal{G}_0(1,2) = 0.7005$$

What is the (patient-average) probability of each arm being allocated in the next block using the GI (and given the priors)?

$$\pi_{1,0} = \frac{(0 \times 1) + (0 \times 1/2 + 1/2 \times 1/2)}{2} = 1/8 \ , \ \pi_{1,1} = \frac{(1 \times 1) + (1 \times 1/2 + 1/2 \times 1/2)}{2} = 7/8.$$

The Forward Looking Gittins Index

Introducing Randomization to the Gittins Index Rule

Assume that T patients are arrive sequentially in blocks of size b over J stages, so that $J \times b = T$. In Villar et al (2015b) we defined group allocation probabilities based on the Gittins Index (GI) rule as follows:

Simplest example: $b = 2$. Priors: control $(s_{(0,0)}, f_{(0,0)}) = (1, 2)$ and experimental $(s_{(1,0)}, f_{(1,0)}) = (1, 1)$

$$j = 1,$$
$$\mathcal{G}_1(1,1) = 0.8699$$
$$\mathcal{G}_0(1,2) = 0.7005$$

$\frac{1}{2}$ / $Y_{1,0} = 1$

$$j = 2,$$
$$\mathcal{G}_1(2,1) = 0.9102$$
$$\mathcal{G}_0(1,2) = 0.7005$$

$\frac{1}{2}$ / $Y_{1,0} = 0$

$$j = 2,$$
$$\mathcal{G}_1(1,2) = 0.7005$$
$$\mathcal{G}_0(1,2) = 0.7005$$

What is the (patient-average) probability of each arm being allocated in the next block using the GI (and given the priors)?

$$\pi_{1,0} = \frac{(0 \times 1) + (0 \times 1/2 + 1/2 \times 1/2)}{2} = 1/8 \ , \ \pi_{1,1} = \frac{(1 \times 1) + (1 \times 1/2 + 1/2 \times 1/2)}{2} = 7/8.$$

FLGI Probabilities: Computation & Properties

A Non-myopic Group Randomised Procedure

C Just as for the MABP, the computational cost of the exact FLGI probabilities grows with the number of arms (K) and b (block size).

Computation in practice can be done via Monte Carlo simulation. Example: $\boldsymbol{P} = [1\ 1\ ;\ 2\ 1\ ;\ 1\ 2\ ;\ 2\ 2]$ ($K = 4$) and block $b = 9$ then $\boldsymbol{\pi} \approx [0.2646\ ;\ 0.5901\ ;\ 0.0246\ ;\ 0.1208]$ after $5 * 10^2$ replicas.

P1 For equal priors the algorithm defines equal allocation probabilities.

P2 As the block size tends to grow (in the limit it equals the trial size), the design tends to a balanced design (given initial equipoise).

P3 If the block is of only 1 patient (i.e. there is an interim after every patient), the FLGI rule recovers the GI rule.

The FLGI in Practice

Example: Redesigning a Real Trial

NeoSphere is a 4-arm ER trial in breast cancer with 417 patients. The response rates reported were 29.0%, 45.8%, 16.8% and 24.0%.

$$H_1 : \boldsymbol{p_1} = [0.29\ 0.458\ 0.168\ 0.24]$$

	Power	Patient Benefit	
	$(1-\beta)$	p^* (s.e.)	*ENS* (s.e.)
ER (block=417)	0.653	0.250 (0.02)	120.88 (9.34)
FLGI (block=9)	0.177	0.804 (0.09)	174.11 (13.3)
GI (block=1)	0.140	0.840 (0.10)	177.97 (13.0)
UB		1	190.99 (0.00)

with the $\pi_{k,j}$ probabilities computed via Monte Carlo simulation.

- Effects of randomisation: (slight) increase in power/ (slight) reduction in ENS (patient benefit) compared to GI
- *Comparable* power levels: apply FLGI to experimental arms only. Allocation to control arm fixed at FR level (25%) (Trippa et al, 2012)

The FLGI in Practice

Example: Redesigning a Real Trial

NeoSphere is a 4-arm ER trial in breast cancer with 417 patients. The response rates reported were 29.0%, 45.8%, 16.8% and 24.0%.

$H_1 : \boldsymbol{p_1} = [0.29\ 0.458\ 0.168\ 0.24]$

	Power	Patient Benefit	
	$(1-\beta)$	p^* (s.e.)	*ENS* (s.e.)
ER (block=417)	0.653	0.250 (0.02)	120.88 (9.34)
C FLGI (block=9)	0.816	0.665 (0.06)	166.40 (11.9)
FLGI (block=9)	0.177	0.804 (0.09)	174.11 (13.3)
GI (block=1)	0.140	0.840 (0.10)	177.97 (13.0)
UB		1	190.99 (0.00)

with the $\pi_{k,j}$ probabilities computed via Monte Carlo simulation.

- Effects of randomisation: (slight) increase in power/ (slight) reduction in ENS (patient benefit) compared to GI
- *Comparable* power levels: apply FLGI to experimental arms only. Allocation to control arm fixed at FR level (25%) (Trippa et al, 2012)

Outline

Incorporating Covariate Information to the Gittins Index

Increasing Patient Benefit by Personalising Treatment

- FDA has recently approved several cancer drugs for use in patients whose tumours have specific genetic characteristics. This has strengthened the promise of "*personalised medicine*"

  But, how can trials answer which types of patients will respond differently to which types of drugs?

  MABP with covariates: let patient outcome $Y_{k,t} \sim Bernoulli(p_k(z_t))$ where $Z_t \sim Bernoulli(q)$ (with q known).
  E.g., $p_k(z_n) = Expit\left(\alpha_k + \beta_k z_t\right)\ \forall t$, where $\text{Expit}(u) = \frac{exp(u)}{1+exp(u)}$.

  For patient t, we observe their covariate value z_t then we treat them.

- Associated MABP with Dynamic Programming: computational complexity even larger than in the classic case. (Deterministic)

Q: Can we define a simple index rule (analogous to GI) in this case? Little work in the literature: Clayton '89; Woodroofe '79

The MABP with covariates and the CARA FLGI

Deriving a covariate-adjusted response-adaptive (CARA) rule

(1) We consider a MABP with K experimental arms, a control arm and T patients. Before arm k is allocated to patient t, a binary covariate Z_t is observed. Immediately after, a binary response $Y_{t,n}$ is observed.

(2) Reformulate the above MABP: for every treatment-covariate combination there exists a combination arm kz. E.g., the arm "00" corresponds to the control arm and covariate negative patients.

New reformulated MABP has $2\,(K+1)$ combinations arms (with rate p_{kt}) and patients are optimally allocated to arms with the constraint that they are only allowed arms feasible given their biomarker profile.

(3) We defined a modified GI rule: each patient gets the treatment with the highest GI among the arms available for their biomarker profile.

(4) From this modified GI, a randomised group allocation procedure is defined as in Villar et al (2015b) but for every covariate value (and block) we have a different vector of allocation probabilities $\pi_{k,j}(Z)$.

The CARA FLGI in Practice

Simulation Results

3-arm trial 300 patients $p_{k0} = (0.22; 0.34; \mathbf{0.49})$, $p_{k,1} = (0.47; \mathbf{0.71}; 0.37)$. Treatment-covariate interaction: best arm for covariate negative patients is arm 2 while for covariate positive patients is arm 1.

	Power		Patient Benefit		
	$(1-\beta_0)$	$(1-\beta_1)$	p_0^* (s.d)	p_1^* (s.d)	ENS (s.d)
ER $_{(b=300)}$	0.82	0.63	0.33 (0.04)	0.33 (0.04)	130.71 (9.3)
CARA C FLGI $_{(b=10)}$	**0.85**	**0.79**	**0.55 (0.16)**	**0.62 (0.06)**	**148.36 (9.6)**
CARA FLGI $_{(b=10)}$	0.13	0.03	0.75 (0.22)	0.86 (0.16)	166.73 (11.2)
CARA GI $_{(b=1)}$	0.11	0.03	0.78 (0.24)	0.88 (0.18)	169.39 (11.4)

CARA FLGI probabilities (Monte Carlo simulation), $T = 300$, $p_z = 0.5$ and 5000 runs.

- Treatment-covariate interactions are detected by the CARA (Covariate-Adjusted Response Adaptive) FLGI procedure but its statistical power is very low.
- In a multi-armed case the CARA CFLGI addresses the power limitation (though in a two-arm setting power may be insufficient).

Outline

Closing Discussion

Opportunities & Remaining Challenges

Armitage (1985)- The search for optimality in clinical trials.

"*I close with two specific suggestions: first, that statisticians concerned with the development of optimization models and those concerned directly in clinical trials should* **meet to discuss** *the feasibility of these methods for various sorts of trials; secondly, that members of the two groups should* **work in collaboration on specific trials** *so as to foster closer understanding and to explore the possibilities in a realistic setting.*"

- Designing implementable optimal designs still requires considerable dialogue between theory and practice. Such a dialogue can potentially lead to sound solutions for the current challenges in clinical trials.

- Explicitly including patient benefit as an optimisation goal can greatly improve trials. Reporting on patient benefit properties of designs should become as standard as reporting expected error rates.

References I

Questions & Comments

Armitage, P. The search for optimality in clinical trials. *International Statistical Review*: 15-24, 1985.

Berry, Donald A Modified two-armed bandit strategies for certain clinical trials *Journal of the American Statistical Association*: 73(362) 339–345, 1978.

Gittins, J. and Jones, D. A dynamic allocation index for the discounted multiarmed bandit problem. *Biometrika*, 66(3):561–565, 1979.

Bather, J.A. On the allocation of treatments in sequential medical trials *International Statistical Review/Revue Internationale de Statistique*,1–13, 1985.

Cheng, Yi and Su, Fusheng and Berry, Donald A Choosing sample size for a clinical trial using decision analysis *Multi-armed bandit allocation indices*, 90(4) 923–936, 2003.

Gittins, J. and Glazebrook, K. and Weber, R. *Multi-armed bandit allocation indices*. Wiley, 2011.

Thompson, W. R. *On the likelihood that one unknown probability exceeds another in view of the evidence of two samples*. Biometrika 25(3/4) 285–294 (1933)

References II

Questions & Comments

Villar, S., Bowden, J. and Wason, J. Multi-armed Bandit Models for the Optimal Design of Clinical Trials: Benefits and Challenges. *Statistical Science Vol. 30, No. 2, 199-215*, 2015.

Villar, S., Wason, J. and Bowden, J. Response-adaptive Randomization for Multi-arm Clinical Trials using the Forward Looking Gittins Index rule *Biometrics Vol. 71, No 4. 969–978*, 2015.

Trippa, L, Lee, E. Wen, P et al *Bayesian Adaptive Randomized Trial Design for Patients with Recurrent Glioblastoma* Journal of Clinical Oncology 30, pp. 3258-3263, 2012.

Villar, S. and Rosenberger, W. Covariate-Adjusted Response-Adaptive Randomization for Multi-Arm Clinical Trials Using a Modified Forward Looking Gittins Index Rule *Biometrics. 2018 Vol. 74, No 1:49–57*, 2018

Rosenberger et al (2001) Covariate-adjusted response-adaptive designs for binary response *Journal of biopharmaceutical statistics* 11(4) 227–236.

Earn-learn dilemma and block size

How to select block size? Should we ramp up accrual?