Optimized Learning While Doing: The REMAP-CAP Adaptive Platform Trial

Derek C. Angus, MD, MPH
Learning While Doing

• Must do two things simultaneously
  • Do: Treat patients as well as possible
  • Learn: Find out what therapies help
Learning While Doing

• Must do two things simultaneously
  • Do: Treat patients as well as possible
  • Learn: Find out what therapies help

• Framed as a (potentially false) choice
Learning While Doing

• Must do two things simultaneously
  • Do: Treat patients as well as possible
  • Learn: Find out what therapies help

• Framed as a (potentially false) choice

• Classic dilemma in decision-making under uncertainty
  • The ‘exploration/exploitation trade-off’
    • James March, Org Sci 1991

• The (elusive) solution is an integrated approach
  • Find the optimal balance to treat patients as well as possible and learn as fast as possible
Outside medicine ...

• Exploration/exploitation (or ‘Learning While Doing’) is everywhere ...
  • Cornerstone of decision-making under uncertainty

• Complex Adaptive Systems research in multiple disciplines
  • Organization science, mathematics, evolutionary biology, economics, social sciences

• Artificial intelligence
  • Reinforcement learning
  • Multi-arm bandits, Markov decision processes, policy evaluations, etc.

• All disciplines exploring the optimal trade-off ...
Inside medicine ...

• ‘Doing’ (practice) and ‘Learning’ (research) are separate
  • Many reasons, including Belmont Report
  • Separate organizations, cultures, people, funding, procedures, and goals

• Consequence: no one really empowered to find the optimal trade-off
  • Always true, but particularly obvious during a pandemic
Best learning tool is the RCT, but 3 major challenges in a pandemic ...

- Randomization is very uncomfortable
  - Physician feels responsible for patient outcomes, consequences are immediate
  - Physician feels less responsible for research, consequences are remote

- RCTs are very cumbersome
  - Slow to start
  - Intrusive to execute

- Little coordination in the clinical research enterprise
  - >100 RCTs registered for HCQ; few likely to be completed
  - AMCs bombarded with 100s of requests to participate in trials; no national or global prioritization
3 solutions from the clinical research enterprise, designed to ‘lean in’ to the realities of clinical care ...

• Make randomization more comfortable
  • Multiple arms, only one is control
  • Adaptive randomization, preferentially assign to best therapy over time

• Make entry into clinical trials ‘1-stop shopping’
  • Simplify interface between clinical practice and clinical research
  • Use master protocols with standard entry criteria, outcomes, etc.
  • Essentially, combine trials/study questions

• Sacrifice ‘sacred cows’ of research
  • Don’t let perfection be the enemy of the good
  • Ex. placebo probably overrated in a pandemic; added rigor not worth the burden
REMAP-CAP Executive Summary

• A global adaptive platform trial

• Designed to determine best treatment for severe pneumonia
  • Randomizes multiple interventions simultaneously, nested within domains
  • Uses a multifactorial Bayesian inference model
  • Uses response-adaptive randomization

• Assesses both interpandemic AND pandemic forms of pneumonia
  • Pre-set rules to switch into pandemic mode

• Entered pandemic mode (termed ‘REMAP-COVID’) in February 2020
Adaptive Platform Trials

• Typically, have focused on pre-approval space
  • Emphasis on efficiency with (very) small sample sizes
  • Different therapies ‘graduate’ to next phase while trial continues

Woodcock and Lavange. *NEJM* 2017
Adaptive Platform Trials Coalition. *Nature Drug Discovery* 2019
New Breast Cancer Results Illustrate Promise and Potential of I-SPY 2 Trial

Trial Identifies Breast Cancer Patients Likely to Benefit from Experimental Drug

By Elizabeth Fernandez on April 07, 2014

Physicians

I-SPY 2 is a collaborative research effort that uses genetic and biological markers from individual patient’s tumors to screen several promising new treatments simultaneously and allows doctors to quickly measure the effectiveness of the treatment prior to removing the tumor.

I-SPY 2 is an Innovative public-private collaboration that combines Personalized Medicine & Novel Trial Design to develop new cancer treatments much faster and for much less cost
Response-adaptive randomization

New patient enrolled; biomarker subtype assessed

Randomly assign to experimental group or control group

For each experimental group, determine adaptive randomization probability within each subtype

Update patient outcome data

Update and apply longitudinal model

Update predictive probability for each experimental regimen vs. control in phase 3 trial for each biomarker signature

Termination rule per group

Stop for futility

Continue

Move to next phase of trial (i.e., “graduate”)

Add new experimental groups if enrollment permits

Update probability in each experimental group vs. control for each subtype

Rugo et al. NEJM 2016
The traditional RCT ...

At the start, 50% chance that A > B
The traditional RCT ...

At the end, >99% sure that A > B

What about in the middle?
A planned trial of A vs. B in 400 patients

The probability that A > B = 78%

Start randomizing MORE patients to A than B ...

After 40 enrolled ...

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alive</td>
<td>14</td>
<td>9</td>
</tr>
<tr>
<td>Dead</td>
<td>6</td>
<td>11</td>
</tr>
</tbody>
</table>
After 80 patients ...

Now, the probability that A > B = 99.9%
Stop the trial!
Caveats

• If the ‘second’ 40 was flat or opposite direction ...
  • Trial continues and the next ‘bet’ swings back closer to 50:50

• When 2 groups, power driven by the smaller group

• So, NOT very helpful if ...
  • Single homogenous cohort
  • Two arms

• But, becomes VERY interesting when ...
  • Multiple arms
  • Multiple subgroups
Response-adaptive randomization

Diagram showing the process of randomization and adaptive randomization in statistical models, with a focus on the interplay between randomization rules and statistical models to generate data.
Response-adaptive randomization

- Odds weighted towards best RX
- Randomization rule
- Statistical model
Response-adaptive randomization

New arms activated

Randomization rule
Statistical model
Response-adaptive randomization

Randomization rule

Statistical model

Or dropped
Response-adaptive randomization

Different weights for different patient groups

Randomization rule

Statistical model
**PLATFORM**
- Perpetual enrollment; continuous learning

**EMBEDDED**
- Align with care; leverage the EHR

**RANDOMIZED**
- Allow CAUSAL inference

**MULTIFACTORIAL**
- Multiple treatments and subgroups

**ADAPTIVE**
- Match odds of success to odds of assignment

**PLATFORM**
- Perpetual enrollment; continuous learning

Angus DC. JAMA 2015
Scenario: 2 of 8 regimens are best

'True' mortality
Scenario: 2 of 8 regimens are best

For a 2,000 patient trial ...

Similar power but 80 fewer deaths

Average results from 1,000s of simulations

‘True’ mortality
REMAP designs ...

• Smart
  • Consider many different treatment options
  • Vary the options depending on the patient

• Safe
  • Probably ‘play’ what is probably the ‘winner’
  • On average, safer ‘in’ the trial than out of it …
REMAP-COVID, a ‘sub-platform’ of REMAP-CAP

• Expanded to all hospitalized patients with COVID-19, in 2 strata
  • Moderate (hospitalized but not severe)
  • Severe (requiring ICU care for respiratory failure or shock)
REMAP-COVID, a ‘sub-platform’ of REMAP-CAP

- Expanded to all hospitalized patients with COVID-19, in 2 strata
  - Moderate (hospitalized but not severe)
  - Severe (requiring ICU care for respiratory failure or shock)

- $1^\circ$ endpoint: organ failure-free days
  - Death worst outcome, followed by number of days free of ICU-based cardiovascular or respiratory support through 21 days
  - Modeled with cumulative logistic proportional odds model

$$
\log \left( \frac{\pi_y}{1 - \pi_y} \right) = [Site] + [Time] + [Age] + \sum_{i=1}^{k} [Intervention] + \sum [Ixl\ Interactions]
$$

- $2^\circ$ endpoints: mortality, WHO ordinal scale, safety
REMAP elements

• Domain – an area where a question is asked ...
  • Domain #1 – choice of antibiotic
  • Domain #2 – whether to give steroids or not
  • Domain #4 – choice of ventilator strategy
  • Etc. ....

• Intervention
  • Any option within a domain ...

• Regimen
  • Unique combination of interventions within a domain ...

• Stratum
  • Baseline subgroup
  • Ex. Moderate vs. Severe COVID19 at presentation
Multifactorial intervention assignments

Regimen = set of domain-specific interventions
Effect of an intervention is conditional upon
• Stratum
• Interventions within other domains

<table>
<thead>
<tr>
<th>Regimen</th>
<th>Domain A</th>
<th>Domain B</th>
<th>Domain C</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1</td>
<td>A1</td>
<td>B1</td>
<td>C1</td>
</tr>
<tr>
<td>#2</td>
<td>A1</td>
<td>B1</td>
<td>C2</td>
</tr>
<tr>
<td>#3</td>
<td>A1</td>
<td>B2</td>
<td>C1</td>
</tr>
<tr>
<td>#4</td>
<td>A1</td>
<td>B2</td>
<td>C2</td>
</tr>
<tr>
<td>#5</td>
<td>A2</td>
<td>B1</td>
<td>C1</td>
</tr>
<tr>
<td>.....</td>
<td>.....</td>
<td>.....</td>
<td>.....</td>
</tr>
<tr>
<td>#n</td>
<td>An</td>
<td>Bn</td>
<td>Cn</td>
</tr>
</tbody>
</table>
REMAP-COVID domains/interventions

• Current COVID19-specific domains
  • Antiviral agents (NONE, HCQ, kaletra, HCQ/kaletra combo)
  • Corticosteroids (NONE, 3 doses)
  • Targeted innate immune modulation (NONE, IL1ra, 2 X IL6ra, IFNbeta, others)
  • Immunoglobulin therapy (NONE, CP, with synthetic IGs to be added later)

• Additional funded domains about to launch
  • Coagulation modulation (prophylaxis only, heparin, possibly dipyridamole)
  • High dose vitamin C (NONE, vitamin C)
  • Statin (NONE, simvastatin)

• Once these 7 domains all running, there are 1,280 separate regimens (recipes) ...
  • Plus, more under development
    • ACE2 modulation (3 subdomains for binding and downstream activation)
    • Ventilation
What does background care look like?

• Surviving Sepsis Campaign Guidelines for COVID19
  • 54 separate care statements
    • Uncertainty regarding every statement

• Even if there are only 2 choices for each of these 54 statements ...
  • $2^{54}$ care ‘regimens’

• In other words, all RCTs are taking place on potentially mammoth scale of background variation in care
REMAP-COVID design

**Background**
COVID emerged in 2019, and its spread was rapid. Several treatments were proposed, but most were untested and unproven. A global collaboration began to address this challenge.

**REMAP-COVID**
- **Main purpose**: To test untested interventions and antigens in real-world patients experiencing COVID-19.
- **First phase**: Six treatment interventions tested in 2020.
- **Current phase**: Four domains: Clinical and EHR embedding, consent and documentation, developing and testing antiviral diagnostic tests, and evaluating and testing biological therapy.

**Current Design**
- **Randomized interventions**
  - Tied to clinical ‘point-of-care’
- **Clinical and EHR embedding**
  - Screen and flag patient
  - Consent documentation
  - Generate regimen order set
  - Flag downstream states
  - Data collection

- **Multifactorial intervention assignments**
  - Regimen = set of domain-specific interventions
  - Effect of an intervention is conditional upon:
    - Stratum
    - Interventions within other domains
  - Domain A, B, C
  - [Table showing assignments]

- **Data collection**
  - Collected at sites
  - Managed at regional data centers
  - Merged at central statistical center

- **Statistical trigger**
  - Result declared when, within stratum, an intervention is:
    - Superior: >95% likely to be best
    - Equivalent: >90% likely that odds within 0.2
    - Inferior: <15% likely to be best

- **External adaptations**
  - Steering Committee can:
    - Add strata, domains & interventions
  - DSMB can:
    - Request new external data be incorporated in priors
    - Overrule statistical triggers

- **Embedding**
  - Patient identification and enrollment
  - Tied to clinical ‘point-of-care’
  - Randomized interventions
  - Issued as ‘order set’ regimen
  - Clinical and EHR embedding
  - Screen and flag patient
  - Consent documentation
  - Generate regimen order set
  - Flag downstream states
  - Data collection

- **Pre-trial design and construction**
  - Pre-specified architecture determined by:
    - Choice of domains, strata, etc.
    - Choice of potential interactions
    - Choices inform a Bayesian inference model
    - Pre-trial simulations evaluate performance
  - Each external adaptation (ex. new domain):
    - Modify elements in Bayesian model
    - Re-simulate before ‘live’ deployment

- **Update and adapt**
  - Re-estimate Bayesian inference model with new data to update probabilities
Regimens, domains and interventions

- Many domains can be added
- ~4 interventions can be tested within any 1 domain @ 1 time
- Interventions can be tested as a ‘nest’
  - Ex. all IL-6 blocking agents vs. none
- A priori consideration re: interactions
- Each domain has a control arm
- If usual care inferior, can be dropped
  - Ex. if all IL6 blockers superior to ‘none’
REMAP-CAP/COVID is global

• A federation of several highly successful clinical trial networks and coordinating centers
  • >100 sites and 13 countries ‘live’
  • New COVID-specific grants from EU, the Netherlands, France, Germany, UK, Ireland, Canada, Australia and NZ

• Scaling up rapidly across the world
  • Funded to expand to >200 sites this month
  • Adding sites in Middle East and South America
  • Discussions for further expansion in Asia (e.g., Japan) and Africa

• Advantage – global positioning allows capture of patients across the globe
Simulations and power

• For ‘head-to-head’ within stratum with no interactions
  • ~400 per group for moderate (OR: 1.7) treatment effect
Ok, but ...

- EHR data quality
- Institutional commitment
- Ethics
- Statistics and design
- Reporting and dissemination of results
- Funding
- Oversight
- Integration with other clinical research programs
A comment on eligibility ...

- Sites can decline to participate in any particular domain or intervention
- Eligibility can also ‘blink’ (temporary inavailability)
- Patients can be ineligible for any particular intervention or domain
- Both patient and site eligibility, by time, is tracked in the model
  - ‘Controls’ are only those who ‘could’ have received an intervention ...
A comment on RAR and contemporaneousness of controls ...

• Principally, patients who receive a given intervention are compared to patients who contemporaneously serve as controls

• But, relative proportions change over time ...
  • Time (by month) included in the model
A comment on suitability for registration ...

• Conceptually, the trial platform is simply ‘hosting’ multiple parallel questions
  • Comparative effectiveness questions
  • Registration trial questions

• Any single domain can run as a free-standing question ...

• Thus, if necessary for a registration trial ...
  • Alpha error control can be specified
  • Placebo (or combination of placebo) can be specified
  • Bounds on RAR can be specified
  • Limits on select ‘co-randomization’ can be specified with other domains
Conclusions

• This pandemic forces us to do 2 things simultaneously
  • Do
  • Learn

• These activities are intertwined: we must ‘Learn While Doing’

• Unfortunately, ‘practice’ traditionally separated from ‘research’
  • The two enterprises must lean in to each other
  • Use ‘learning designs’ that accommodate ‘doing’ at the same time

• Global adaptive platform trials have potential as LWD instruments