# Innovative Trial Designs: A Triple Win for Patients, Health Care Providers, and Manufacturers

Jason Connor[1,2], Jordan Elm[3], Kristine Broglio[1], & Jaideep Kapur[4]
[1]Berry Consultants, [2]University of Central Florida College of Medicine,
[3]Medical University of South Carolina, [4]University of Virginia

2013 M-CERSI
University of Maryland
September 5, 2013

# Conflicts of Interest

- None for three drugs being discussed

- Consult with many other drug & device companies I won't discuss

# Acknowledgements


## ADAPT-IT U01-NS073476

- William Barsan, MD
  University of Michigan
- Donald Berry, PhD, Scott Berry, PhD
  MD Anderson, Berry Consultants
- Roger Lewis, MD, PhD
  UCLA Harbor
- Valerie Durkalski, PhD, Yuko Palesch, PhD
  Medical University of South Carolina
- Michael Fetters, MD, Laurie Logocki, PhD
  University of Michigan
- Shirley Frederickson, RN, MS
  University of Michigan
- Will Meurer, MD
  University of Michigan
- Robin Conwit, MD & Scott Janis, MD
  NINDS
- Kristine Broglio, MS
  Berry Consultants

## ESETT

- Jaideep Kapur, MD
  University of Virginia
- Jordan Elm, PhD
  Medical University of South Carolina
- James Chamberlain, MD
  Children's National Med Center
- Nathan Fountain, MD
  University of Virginia
- Daniel Lowenstein, MD
  UCSF
- Shlomo Shinnar, MD, PhD
  Albert Einstein COM
- Rob Silbergleit, MD
  University of Michigan
- David Treiman, MD
  Barrow Neurological Institute
- Wenle Zhou, PhD
  Medical University South Carolina

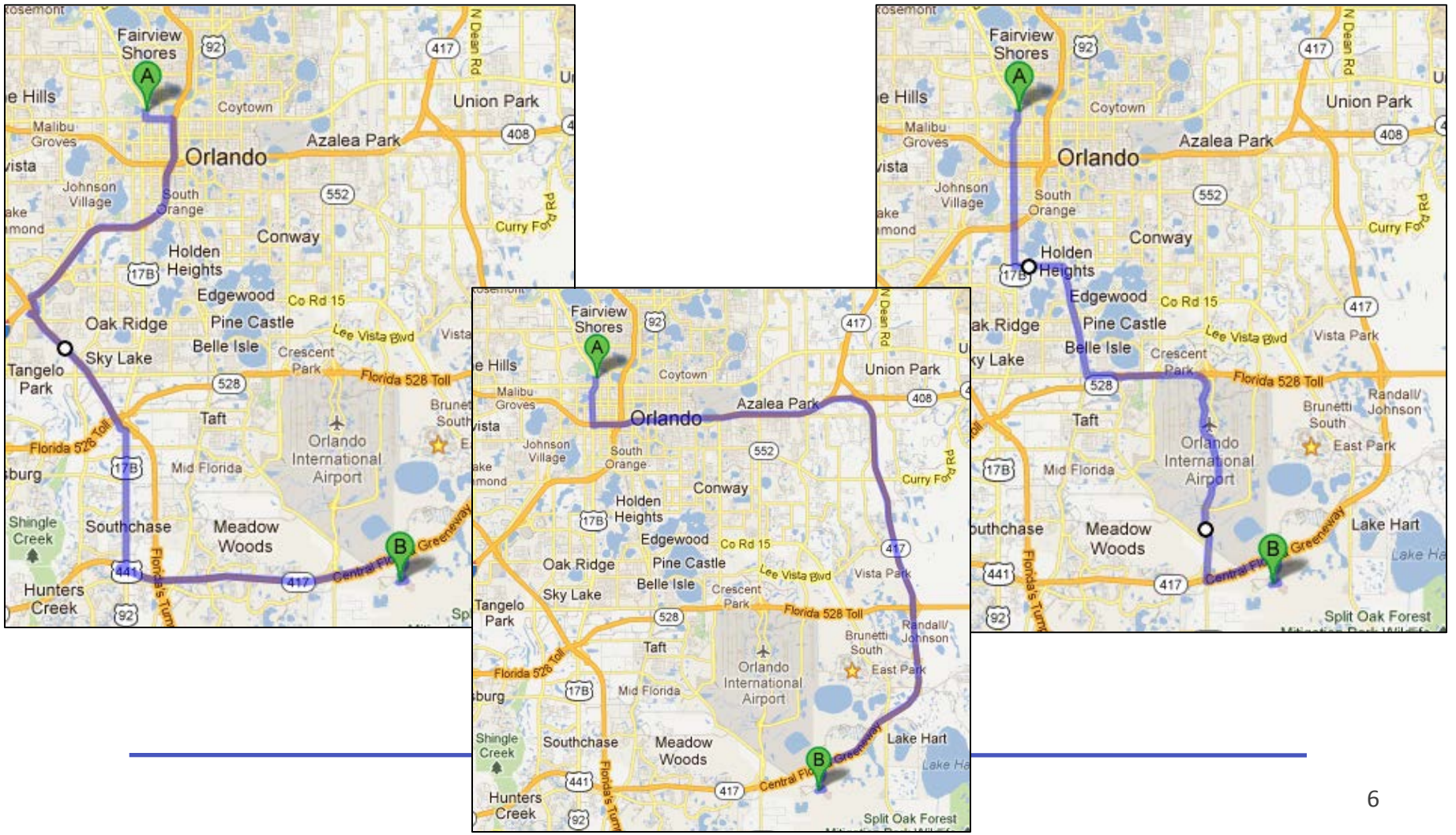
# Research Question

- How to treat static epilepticus patients who've failed benzodiazapines?
  - fosphenytoin (fPHT)
  - levetiracetam (LVT)
  - valproic acid (VPA)
- Composite endpoint
  - Cessation of seizure with 20 minutes
  - No further treatment within 2 hours
  - No significant adverse event

# Comparative Effectiveness


- No control group
  - Three drugs start out equal
  - Want to know which is best
- What is Type I error in CER?
  - Consequence of Type I error less in CER
- Really want to know
  - Which drug is best … with measure of certainty
  - Which drug is worst … with measure of certainty

# A Common CER Trial



# Bayesian Adaptive Design Features


- Adaptively allocate to favor better treatments
- Drop poor performing arms
  - Relative to one another
  - Relative to 25% goal
- Stop early if we know the answer
  or know we won't know
  - Efficacy stop if treatment clearly better
  - Futility stop if unlikely to ID a 'best' or 'worst'
    - Do not stop if 1 worse and other 2 equally good
  - Futility stopping if all arms bad

# Adaptive Allocation


- Randomize 300 patients equally
- At 300 & then every 100 adaptively allocate to
  - Favor better performing treatments
  - Favor treatments with greater uncertainty
  - Every 100 = About every 6 months | expected accrual

$$r_t \propto \sqrt{\frac{\Pr\left(p_t = \max(p)\right) Var(p_t)}{n_t}}$$

- If allocation probability < 5%, suspend accrual
- If Pr(Success > 0.25) < 0.05 drop arm

# Early Stopping

- Begins after 400 patients and every additional 100 patients accrued
- Early Success Stopping:
  - If arm has 97.5% probability of having highest success rate
    - i.e. Pr(Arm = Max Effective Trt) > 0.975
- Early Futility Stopping
  - If predicted probability of success (ID ‘winner’ or ‘loser’ at the max N=795) < 0.05
  - If all doses have Pr(Success > 0.25) < 0.05

# Example Trial: 300-pt analysis

| | N Enrolled<br>Observed Response Rate | | | Pr(Max Effective Trt) | | | Pr(Allocation) | | | Pred Prob |
|---|---|---|---|---|---|---|---|---|---|---|
| Look | LVT | fPHT | VPA | LVT | fPHT | VPA | LVT | fPHT | VPA | |
| 300 | 51/100<br>51% | 55/100<br>55% | 64/100<br>64% | 0.025 | 0.092 | 0.88 | 0.12 | 0.22 | 0.66 | 0.71 |
| | | | | | | | | | | |
| | | | | | | | | | | |
| | | | | | | | | | | |
| | | | | | | | | | | |

# Example Trial: 400-pt analysis

| | N Enrolled<br>Observed Response Rate | | | Pr(Max Effective Trt) | | | Pr(Allocation) | | | Pred Prob |
|---|---|---|---|---|---|---|---|---|---|---|
| Look | LVT | fPHT | VPA | LVT | fPHT | VPA | LVT | fPHT | VPA | |
| 300 | 51/100<br>51% | 55/100<br>55% | 64/100<br>64% | 0.025 | 0.092 | 0.88 | 0.12 | 0.22 | 0.66 | 0.71 |
| Next 100 | 6/11<br>55% | 19/26<br>73% | 39/63<br>62% | | | | | | | |
| 400 | 57/111<br>51% | 74/126<br>59% | 105/163<br>64% | 0.01 | 0.16 | 0.83 | 0.09 | 0.34 | 0.57 | 0.50 |
| | | | | | | | | | | |
| | | | | | | | | | | |

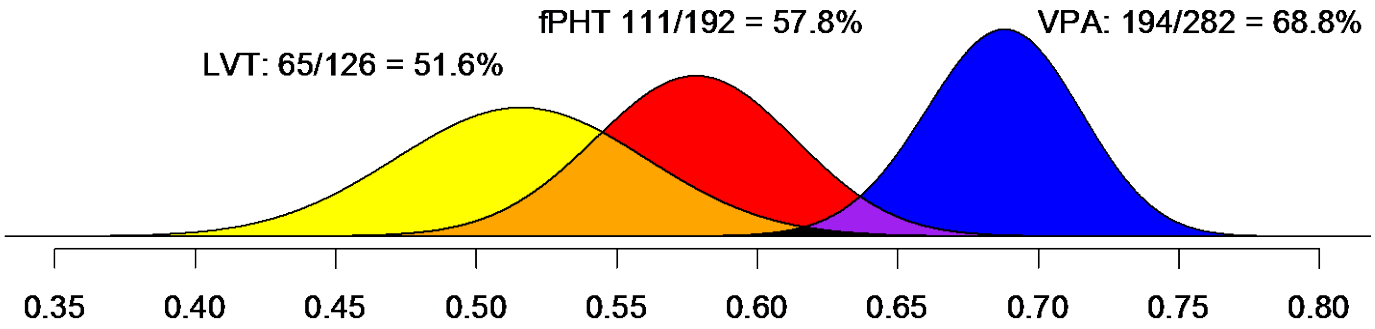
# Example Trial: 500-pt analysis

| | N Enrolled<br>Observed Response Rate | | | Pr(Max Effective Trt) | | | Pr(Allocation) | | | Pred Prob |
|---|---|---|---|---|---|---|---|---|---|---|
| Look | LVT | fPHT | VPA | LVT | fPHT | VPA | LVT | fPHT | VPA | |
| 300 | 51/100<br>51% | 55/100<br>55% | 64/100<br>64% | 0.025 | 0.092 | 0.88 | 0.12 | 0.22 | 0.66 | 0.71 |
| 400 | 57/111<br>51% | 74/126<br>59% | 105/163<br>64% | 0.01 | 0.16 | 0.83 | 0.09 | 0.34 | 0.57 | 0.50 |
| Next 100 | 5/12<br>42% | 20/38<br>53% | 34/50<br>68% | | | | | | | |
| 500 | 62/123<br>50% | 94/164<br>57% | 139/213<br>65% | 0.004 | 0.056 | 0.94 | 0.08 | 0.23 | 0.69 | 0.59 |

# Example Trial: 600-pt analysis

| | N Enrolled<br>Observed Response Rate | | | Pr(Max Effective Trt) | | | Pr(Allocation) | | | Pred Prob |
|---|---|---|---|---|---|---|---|---|---|---|
| Look | LVT | fPHT | VPA | LVT | fPHT | VPA | LVT | fPHT | VPA | |
| 300 | 51/100<br>51% | 55/100<br>55% | 64/100<br>64% | 0.025 | 0.092 | 0.88 | 0.12 | 0.22 | 0.66 | 0.71 |
| 400 | 57/111<br>51% | 74/126<br>59% | 105/163<br>64% | 0.01 | 0.16 | 0.83 | 0.09 | 0.34 | 0.57 | 0.50 |
| 500 | 62/123<br>50% | 94/164<br>57% | 139/213<br>65% | 0.004 | 0.056 | 0.94 | 0.08 | 0.23 | 0.69 | 0.59 |
| Next 100 | 3/3<br>100% | 17/28<br>61% | 55/69<br>80% | | | | | | | |
| 600 | 65/126<br>52% | 111/192<br>58% | 194/282<br>69% | 0.000<br>0.87 | 0.008<br>0.13 | **0.992**<br>0.00 | Trial Stops Early for Identifying Best Treatment | | | |

# Example Trial: Final Evaluation




| Treatment | Observed | % | 95% CI | Pr(Best) | Pr(Worst) |
|---|---|---|---|---|---|
| VPA | 194/282 | 68.8% | (.632, .739) | 0.992 | 0.0005 |
| fPHT | 111/192 | 57.8% | (.507, .646) | 0.007 | 0.138 |
| LVT | 65/126 | 51.6% | (.429, .601) | 0.0005 | 0.862 |

| Difference | Observed | 95% CI | Pairwise Comparison |
|---|---|---|---|
| VPA – fPHT | 0.110 | (0.022, 0.197) | Pr(VPA>fPHT) = 0.993 |
| VPA – LVT | 0.172 | (0.069, 0.272) | Pr(VPA>LVT) > 0.999 |
| fPHT - LVT | 0.062 | (-0.049, 0.172) | Pr(fPHT>LVT) = 0.862 |

# Comparison to without Adaptive Randomization

| | Adaptive Randomization | | | Fixed Randomization | | |
|---|---|---|---|---|---|---|
| **Scenario 3 Efficacy Rates** | **Power Best/Wst** | **Mean N** | **% to Best** | **Power Best/Wst** | **Mean N** | **% to Best** |
| Null 0.5 – 0.5 – 0.5 | 0.017 0.021 | 508 | | 0.023 0.009 | 499 | |
| One Good 0.5 – 0.5 – 0.65 | 0.90 0.04 | 484 | 47 | 0.88 0.04 | 492 | 33 |
| Two Good 0.5 – 0.65 – 0.65 | 0.11 0.69 | 684 | 83 | 0.10 0.80 | 686 | 67 |
| One Middle One Good 0.5 – 0.575 – 0.65 | 0.49 0.24 | 592 | 47 | 0.45 0.31 | 597 | 33 |
| All Bad 0.25– 0.25 – 0.25 | 0.016 0.020 | 522 | | 0.021 0.009 | 507 | |
| All Very Bad 0.10 – 0.10 – 0.10 | 0.005 0.000 | 400 | | 0.007 0.000 | 400 | |

# Comparison to without Adaptive Randomization

| | Adaptive Randomization | | | Fixed Randomization | | |
|---|---|---|---|---|---|---|
| **Scenario 3 Efficacy Rates** | **Power Best/Wst** | **Mean N** | **% to Best** | **Power Best/Wst** | **Mean N** | **% to Best** |
| Null 0.5 – 0.5 – 0.5 | 0.017 0.021 | 508 | | 0.023 0.009 | 499 | |
| One Good 0.5 – 0.5 – 0.65 | 0.90 0.04 | 484 | 47 | 0.88 0.04 | 492 | 33 |
| Two Good 0.5 – 0.65 – 0.65 | 0.11 0.69 | 684 | 83 | 0.10 0.80 | 686 | 67 |
| One Middle One Good 0.5 – 0.575 – 0.65 | 0.49 0.24 | 592 | 47 | 0.45 0.31 | 597 | 33 |
| All Bad 0.25– 0.25 – 0.25 | 0.016 0.020 | 522 | | 0.021 0.009 | 507 | |
| All Very Bad 0.10 – 0.10 – 0.10 | 0.005 0.000 | 400 | | 0.007 0.000 | 400 | |

# Comparison to without Adaptive Randomization

| | Adaptive Randomization | | | Fixed Randomization | | |
|---|---|---|---|---|---|---|
| **Scenario 3 Efficacy Rates** | **Power Best/Wst** | **Mean N** | **% to Best** | **Power Best/Wst** | **Mean N** | **% to Best** |
| Null 0.5 – 0.5 – 0.5 | 0.017 0.021 | 508 | | 0.023 0.009 | 499 | |
| One Good 0.5 – 0.5 – 0.65 | 0.90 0.04 | 484 | 47 | 0.88 0.04 | 492 | 33 |
| Two Good 0.5 – 0.65 – 0.65 | 0.11 0.69 | 684 | 83 | 0.10 0.80 | 686 | 67 |
| One Middle One Good 0.5 – 0.575 – 0.65 | 0.49 0.24 | 592 | 47 | 0.45 0.31 | 597 | 33 |
| All Bad 0.25– 0.25 – 0.25 | 0.016 0.020 | 522 | | 0.021 0.009 | 507 | |
| All Very Bad 0.10 – 0.10 – 0.10 | 0.005 0.000 | 400 | | 0.007 0.000 | 400 | |

# Comparison to without Adaptive Randomization

| | Adaptive Randomization | | | Fixed Randomization | | |
|---|---|---|---|---|---|---|
| **Scenario 3 Efficacy Rates** | **Power Best/Wst** | **Mean N** | **% to Best** | **Power Best/Wst** | **Mean N** | **% to Best** |
| Null 0.5 – 0.5 – 0.5 | 0.017 0.021 | 508 | | 0.023 0.009 | 499 | |
| One Good 0.5 – 0.5 – 0.65 | 0.90 0.04 | 484 | 47 | 0.88 0.04 | 492 | 33 |
| Two Good 0.5 – 0.65 – 0.65 | 0.11 0.69 | 684 | 83 | 0.10 0.80 | 686 | 67 |
| One Middle One Good 0.5 – 0.575 – 0.65 | 0.49 0.24 | 592 | 47 | 0.45 0.31 | 597 | 33 |
| All Bad 0.25– 0.25 – 0.25 | 0.016 0.020 | 522 | | 0.021 0.009 | 507 | |
| All Very Bad 0.10 – 0.10 – 0.10 | 0.005 0.000 | 400 | | 0.007 0.000 | 400 | |

# Summary

- Comparative Effectiveness is an ideal setting for Bayesian adaptive trials
  - Higher power
  - Smaller sample size
  - Treat more patients with best therapy
- Interim analyses early & often
- Steer patients to better therapies
- Steer patients away from poorer therapies
- Stop trial as soon as best treatment identified
- Natural way of learning & making decisions