Adaptive group-sequential multi-regional outcome studies in vaccines

Inna Perevozskaya, Statistical Research and Consulting Center, Pfizer

Joseph Severs, Vaccines Biostatistics, Pfizer

Acknowledgements

- Yahong Peng
- Scott Patterson
- Neal Thomas
- Christy Chuang-Stein
- Qin Jiang

Outline

- Motivating example
- Background: statistical methods for vaccine efficacy studies
- Clinical development strategy options:
  1. Adaptive Phase 2/3 design
  2. Simple Group-Sequential Ph2b design followed by pivotal Ph3
  3. Semi-pivotal Ph2b followed by Ph3
- Other design complicating factors:
  - heterogeneity of population
  - multiple regions
- Summary and Conclusions

Motivating Example

- Vaccine efficacy study
- Vaccine is supposed to prevent post-surgical infection
  - Administered prior to surgery
  - Rate of post-operative infection monitored within fixed time interval after surgery
- Without vaccine, the rate of infection is very low
- The disease is very serious and there are no good treatment options available, i.e. unmet medical need
- Such situations encourage accelerated development and approval is sometimes possible based on one pivotal study only
- Time-to market is crucial due to competitive landscape
  - Development has to be completed within tight timeframe
- That motivated the study team to explore adaptive design options for this study/program

Background : statistical approaches to designing vaccine efficacy trials

- In a randomized, placebo controlled trial , vaccine efficacy is defined as

$$VE=1-RR$$

- There are 3 classes of methods, differing in how follow-up is accounted for and how the Relative Risk is defined
  1. RR is ratio of 2 attack rates ($\theta=\pi_{trt} / \pi_{pbo;}$ fixed follow-up time)
  2. RR is ratio of 2 infection rates ($\theta=\pi_{trt} / \pi_{pbo;}$ variable follow-up time, rates are per-person years)
  3. RR is ratio of 2 force of infection rates (time-to event analysis)
- Methods 1 and 2 are very close if surveillance period is fixed and event rate is low

Statistical analysis of vaccine efficacy data

$$VE = (1-\theta) * 100\% = \left(1 - \frac{\pi_1}{\pi_0}\right) * 100\%$$

$$\pi_1 = \frac{s_1}{T_1} \quad \pi_0 = \frac{s_0}{T_0}$$

s_1, s_0 –number of events in treatment and placebo groups

T_1, T_0 – exposures (person−years)

$$s_1 \sim Poi(\lambda_1), \quad s_0 \sim Poi(\lambda_0)$$

Conditional on total number of events $s = s_0 + s_1$,

$$s_1 \sim Bin(s, \pi)$$

$$\pi = \frac{T_1\lambda_1}{T_1\lambda_1 + T_0\lambda_0} = \{\text{if equal randomization}\} = \frac{\theta}{\theta+1}$$

$$\theta = \frac{\pi}{1-\pi} \quad \text{VE} = \left(1 - \frac{\pi}{1-\pi}\right) * 100\%$$

Re-stating hypothesis testing in terms of binomial proportion

- Clinicians prefer to state hypothesis in terms of VE, with or without super-efficacy requirement
  1. $H_0: VE \leq 0 \; vs. H_A: VE > 0$
  2. $H_0: VE \leq \delta \; vs. H_A: VE > \delta \; (e.g. \delta = 30\%$ for "super efficacy")
- Given relationship between VE and π,

$$\pi = \frac{\theta}{\theta + 1} \; ; \quad \theta = (1 - \mathrm{VE}/100)$$

- Hypothesis testing can be re-written in terms of π:
  1. $H_0: \pi \geq 0.50 \; vs. H_A: \pi < 0.50$
  2. $H_0: \pi \geq 0.41 \; vs. \; H_A: \pi < 0.41$ ("super efficacy")

Power calculations for fixed event design

- In fixed event design, study power calculations are performed using conditional distribution of s_1 (number of events coming from vaccine group) given s (total number of events)
  - Find a pair (s, s_1) such that 95% LCB for VE exceeds δ (super-efficacy threshold) and actual power is ≥80%
  - Power=$Pr(\text{to see } \leq s_1/s \text{ events in vaccine group}|H_A)$
- Power is driven by total number of events *s* accrued; power calculation depends on underlying VE only (not λ_0)
- By conditioning on total number of events, we don't have to worry about attack rate in the placebo group (would be nuisance parameter in a two–sample binomial problem)
- But , in turn our sample size (to accrue s events) becomes a random variable and placebo attack rate λ_0 affects it

Power calculations for fixed event design (cont.)

- $N = 2 * \frac{s}{\lambda_0 * (2 - VE)} -$ expected sample size to accrue s events
- If λ_0 is small, the sample size *N* is large; that's why "rare" diseases pose a challenge
- If VE close to 1, N also grows large for fixed *s*
- Design challenge:
  - both parameters were very uncertain at the design stage
    - placebo rate reported in literature ranged from 0.06% to 3%
    - VE of interest would be 50%-80%
  - both can impact the resulting sample size quite dramatically
- Next slide gives examples to illustrate impact

Impact of VE and placebo attack rate on sample size (α=0.05, power 90%, super-efficacy δ=30%)

Sc#	VE	λ_0	s	s_1	s_0	n_1	n_0
1	0.6	0.03	154	51	103	3667	3667
2	0.6	0.02	154	51	103	5501	5501
3	0.6	0.01	154	51	103	11001	11001
4	0.7	0.03	69	20	49	1770	1770
5	0.7	0.02	69	20	49	2654	2654
6	0.7	0.01	69	20	49	5308	5308
7	0.8	0.03	38	9	29	1056	1056
8	0.8	0.02	38	9	29	1584	1584
9	0.8	0.01	38	9	29	3167	3167

Recap : study design challenge

- A very large multi-center study would be required to demonstrate super efficacy (ultimate goal)
- Very little was know about potential VE and placebo attack rate
  - Placebo attack rate (nuisance parameter λ_0) is highly influential in sample size calculation
  - Unfortunately it is unknown and variable across study sub-populations (surgical subgroups) according to literature
  - Heterogeneity of study population :
    - Primary outcome (infection) could be affected
    - Multiple centers/regions
    - Like any trial where surgical procedure is involved, quality of care may impact primary endpoint (infection)
    - Underlying background infection risk varies with type of surgical procedure and other pre-existing co-morbidities

Initial clinical development thoughts

- Since the proposed study would be the first one to evaluate efficacy and required very large sample size, the study team was asked to "de-risk " it by utilizing an adaptive design
- Initial AD proposal on the table included a complicated option
  - to address multiple sub-populations
  - as well as uncertain treatment effect & event rate
  - all at once in one trial
- After many deliberations it was decided to put the heterogeneity issue aside and focus on one sub-population only
  - where background rate is highest
- That would allow to study vaccine efficacy in "fastest" way possible
  - then augment the population in a separate trial if the results are promising
- Even then such " fast" trial would be huge and long
  - additional de-risking by adding early futility stop and/or sample size re-estimation would be needed

Clinical Development Program Option 1

- No Phase 2b study: Proceed into Ph3 directly from Ph2a
- Add very early futility analysis and possible SSR later on
- Assumptions:
  - True VE=60%,
  - δ=30% for LCB ("super-efficacy" requirement)
  - Power 90% and α=0.05 (2-sided)
  - Underlying effect rate in placebo is λ_0=0.02
- Required sample size for fixed design:
  - ~N=5501 patients per treatment arm needed
  - To obtain S=154 total events
  - primarily driven by low underlying placebo event rate (λ_0=0.02) and super-efficacy requirement δ=30%
- **Decision criteria for study success:**
  - **$\leq S_1$=51 out of these 154 has to come from vaccines group in order for (1-α)*100% LCB to be > δ=30%**

De-risking strategies based on conditional power

- To de-risk this large study, early futility analysis based on CP was proposed after 16 events (~10% information)
- Conditional power is a widely-used concept to make interim decisions during an “adaptive” trial
- It quantifies the likelihood of achieving final study success given observed interim data
- The basic principle is to terminate the study early if conditional power is low (<30%) , proceed without modifications if it is high (>80%) and possibly increase sample size at interim if the probability of final success is in-between (30%-80%)
  - Other adaptations (such as SSR or even population enrichment were considered but at a later interim looks
  - Focus of 1st interim look was futility only

Futility rule based on conditional power

- Conditional Power = Pr (Success at the end of trial | observed data and assumptions about treatment effect)
- Success was defined as observing ≤51/154 events in vaccines group
- Suppose we do one IA for futility only at *s*=16 events
- And observe (s_1, s_0) "split" for vaccine and placebo, respectively
  - CP1=Pr (to see ≤ (51- s_1) / (154-16) events in vaccine | VE=0.6) ← alternative hypothesis
  - CP2=Pr (to see ≤ (51- s_1) / (154-16) events in vaccine | estimated VE) ← observed data
- Note: both rules depend on assumptions +observed interim data but to a different degree:
  - CP1 is more dominated by "belief" that VE is as hypothesized, i.e. VE =0.6
  - CP2 is more sensitive to interim data than CP1
  - Design1: If CP1 < 30 % after 16 events then stop for futility
  - Design2: If CP2 < 30% after 16 events then stop for futility

Illustration of CP1 and CP2 decision rules for total S=16 cases and various split scenarios

total	s1-(vac)	s0-(pbo)	ve-alt	pi-alt	CP1	pi_est	ve_est	CP2
16	0	16	0.6	0.29	0.99	0.00	1.0	1.00
16	1	15	0.6	0.29	0.98	0.06	0.9	1.00
16	2	14	0.6	0.29	0.97	0.13	0.9	1.00
16	3	13	0.6	0.29	0.95	0.19	0.8	1.00
16	4	12	0.6	0.29	0.93	0.25	0.7	0.99
16	5	11	0.6	0.29	0.91	0.31	0.5	0.73
16	6	10	0.6	0.29	0.87	0.38	0.4	0.14
16	7	9	0.6	0.29	0.83	0.44	0.2	0.00
16	8	8	0.6	0.29	0.78	0.50	0.0	0.00
16	9	7	0.6	0.29	0.72	0.56	-0.3	0.00
16	10	6	0.6	0.29	0.66	0.63	-0.7	0.00
16	11	5	0.6	0.29	0.59	0.69	-1.2	0.00
16	12	4	0.6	0.29	0.51	0.75	-2.0	0.00
16	13	3	0.6	0.29	0.44	0.81	-3.3	0.00
16	14	2	0.6	0.29	0.36	0.88	-6.0	0.00
16	15	1	0.6	0.29	0.29	0.94	-14.0	0.00
16	16	0	0.6	0.29	0.23	1.00		0.00

FUTILITY STOP

What does it mean in practice?

- If IA is conducted with S=16 total cases:

- CP1 rule: terminate if 15/16 or 16/16 in vaccine

  Note: this rule “believes” in hypothesis VE=0.6, so the data has to be pretty extreme not in favor of vaccine to trigger an early stop

- CP2 rule: terminate if split is ≥ 6/16 cases in vaccine

  Note : this rule relies on interim estimate of VE => more aggressive stopping

What does it mean in practice? (cont.)

- If IA conducted after s=24 cases:
  - CP1 rule: terminate if ≥ 18/24 in vaccine group
  - CP2 rule: terminate if ≥ 9/24 in vaccine group

- If IA conducted after s=40 cases:
  - CP1 rule: terminate if ≥ 22/40 in vaccine group
  - CP2 rule: terminate if ≥ 15/40 in vaccine group

Operating characteristics of 2 rules an different VE scenarios

- For various VE scenarios, run 10k " trials" and tabulate percentage of times we stop early for futility with each rule:

	True VE=0	True VE=0.1	True VE=0.2	True VE=0.3	True VE=0.4	True VE=0.5	True VE=0.6
S=16, CP1 rule	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S=16, CP2 rule	0.90	0.85	0.79	0.70	0.59	0.46	0.29

- CP2 does much better than CP1 when there is no efficacy
- BUT it's stopping too aggressively in case of a good drug
  - kills a borderline-good drug (VE=0.5) with almost 50% probability
  - Kills a good drug (VE=60%) with ~30% probability
  - this is driven not only by too early of an interim but high bar of 30% superefficacy

Operating characteristics of 2 CP rules, by VE scenario and IA timing

Table 1 Probability of interim stop under various scenarios of IA1 timing and true VE (using CP1 rule based on effect size in alternative hypothesis)

Futility Interim timing (total events)	True VE=0	True VE=0.1	True VE=0.2	True VE=0.3	True VE=0.4	True VE=0.5	True VE=0.6
S=16	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S=24	0.01	0.01	0.00	0.00	0.00	0.00	0.00
S=40 (25% info)	0.32	0.21	0.12	0.05	0.02	0.00	0.00
S=60 (40% info)	0.81	0.70	0.52	0.32	0.15	0.04	0.00

Table 2 Probability of interim stop under various scenarios of IA1 timing and true VE (using CP2 rule based on observed rather than hypothesized effect

Futility Interim timing (total events)	True VE=0	True VE=0.1	True VE=0.2	True VE=0.3	True VE=0.4	True VE=0.5	True VE=0.6
S=16	0.90	0.85	0.79	0.70	0.59	0.46	0.29
S=24	0.93	0.87	0.81	0.71	0.58	0.41	0.22
S=40	0.96	0.92	0.85	0.74	0.57	0.34	0.14
S=60	0.98	0.97	0.91	0.79	0.60	0.34	0.11

Summary of Development option #1

- The option of single Ph3 study with early stop for futility at s=16 events was rejected
  - Futility criteria did not perform well across all possible VE scenarios
- The only way to get “reasonable” performance was to move increase 1st IA timing to 40% of information
- That finally convinced the team that a proper Ph2b study, separate from Ph3 may be needed

Clinical Development Option #2

- Smaller Ph2b study to determine “proof-of-principle”
  - GSD with early futility assessment half-way
- Followed by separate Ph3
  - Either GSD or adaptive SSR

Phase 2b design

- Assumptions:
  - True VE=60%
  - δ=0 for LCB (no "super-efficacy" requirement)
  - Power 80% and α=0.1 (1-sided)
  - Underlying effect rate in placebo is λ_0=0.02
- Required sample size for fixed design (without interim):
  - ~N=1858 patients overall
  - To obtain S=26 total events
- **Decision criteria for (fixed design) study success:**
  - **$\leq S_1$=9 out of these 26 has to come from vaccines group in order for (1-α)*100% LCB to be > δ=0**
- We add interim analysis options to this "default" design and look at its type1/2 errors and other operating characteristics

Phase Ph2b design with interim analysis

- Classical Group-Sequential (GS) methodology which is classified as one of the “well-understood” designs by the FDA Adaptive Design guidance was utilized
  - Design with O’Brien-Fleming (OBF) type of boundary for single binomial sample was used
  - Following the framework described in Jennison & Turnbull (2000) Ch12, used OBF boundaries based on normal approximation as a starting point
  - Power and type 1 error may not hold well in case of small samples (n=26)
  - will “refine” the boundary using exact binomial calculations

Obtaining GSD boundaries from EAST Software

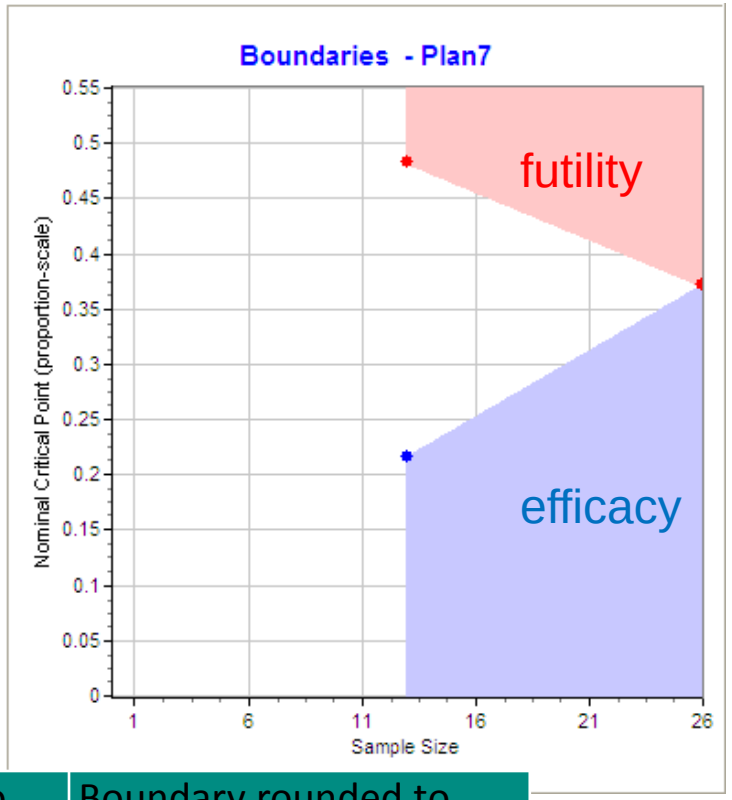
Notes:

- overall sample size S=26 events for a fixed design was obtained via exact calculation based on binomial distribution
- Using 80% power, 60% VE, no super-efficacy claim (δ=0)
- Power of 0.82 given here is based on normal approximation, i.e. not 100 % accurate
- The main goal of this design is to get boundaries (highlighted)

	Plan7
Test type	1-Sided
Alpha	0.1
Power	0.821
# looks incl final	2
spacing of looks	Equal
Eff. or Fut stop	**H0 or H1 (NB)**
OBF boundary for eff.	**LD (OF)**
OBF boundary for fut	**LD (OF)**
π for Null	0.5
π for Alternative	0.285
variance calc.	Under Null
Fixed max ssize	26
Expected ssize under H0	19
Expected ssize under Ha	21
Expected ssize under ½ Ha	22

Obtaining GSD boundaries from EAST Software (cont.)

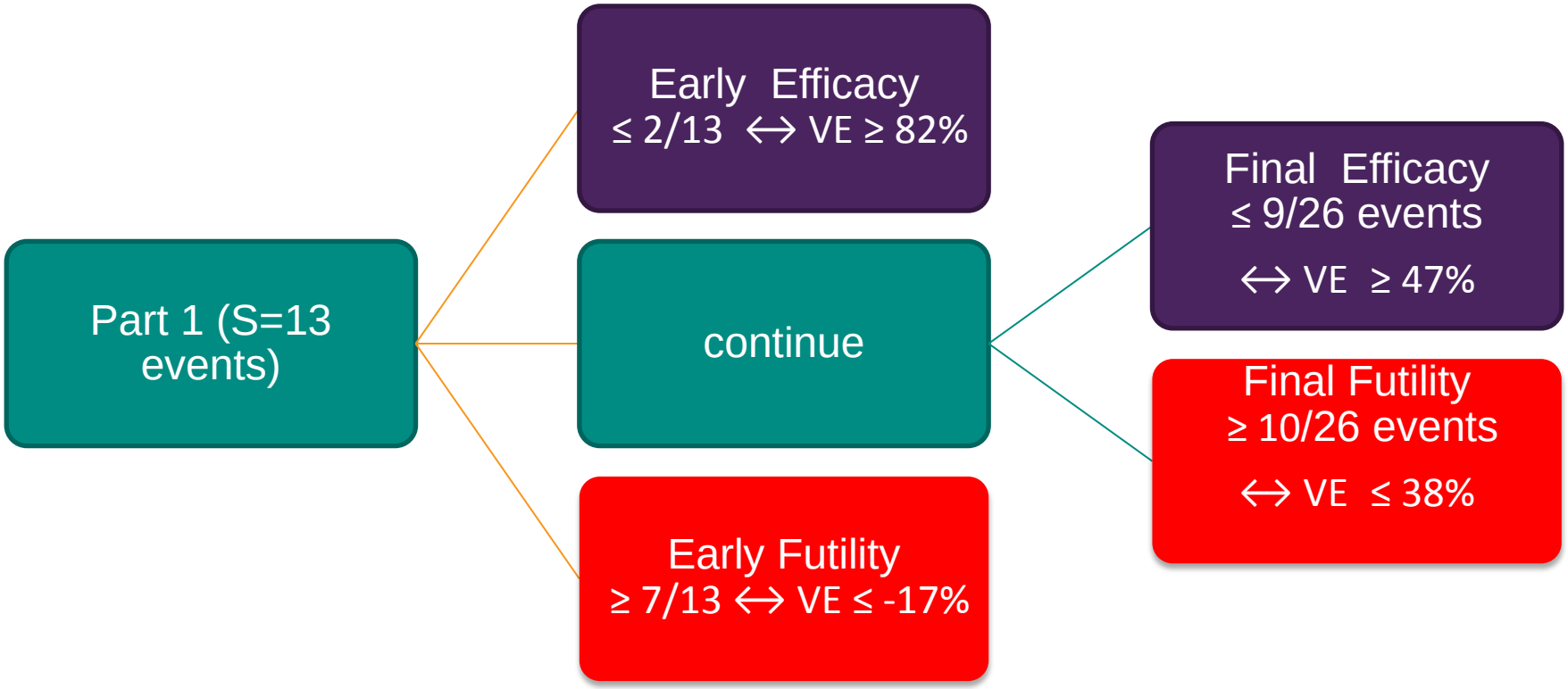
- Boundaries expressed in terms of events rather than proportions
- "discretized boundaries " are passed on to gsdesign to compute operating characteristics based on exact binomial dist's
- *gsbin2<-gsBinomialExact(k=2,theta=c(0.5, 0.285), n.I=c(13,26),a=c(2,9), b=c(7,10))*



	Boundary from EAST: prop. Scale		Boundary converted to event scale		Boundary rounded to integer: gsDesign inputs	
Sample Size	efficacy	futility	efficacy	futility	a (lower)	b (upper)
S=13 (1st IA)	0.215223	0.482315	2.797896	6.270091	**2**	**7**
S=26(final)	0.37091	0.37091	9.64367	9.64367	**9**	**10**

Ph2 GSD with early stopping for efficacy or futility

Note: the final decision rule is same as fixed design due to discrete nature of binomial distribution



Operating characteristics of Ph2 GSD under variety of VE scenarios

	π	Pr(early eff. stop)	Pr(early fut. Stop)	Pr(final success)	Pr (final failure)	Overall Pr. of Success	Ave. # events
VE=0%	0.500	0.011	0.500	0.073	0.415	0.085	19
VE=10%	0.474	0.018	0.424	0.115	0.443	0.133	20
VE=30%	0.412	0.049	0.257	0.267	0.428	0.315	22
VE=50%	0.333	0.139	0.103	0.498	0.260	0.638	23
VE=60%	0.286	0.235	0.049	0.575	0.141	0.810	22
VE=70%	0.231	0.393	0.016	0.545	0.046	0.938	21
VE=80%	0.167	0.627	0.002	0.365	0.006	0.992	18

Type 1 Error

Power

Re-cap of Initial Ph2 GSD proposal (option 2)

- The design was very simple in contrast to the original proposal of complicated single ph3 study
- It also was less ambitious (looser type1/2 error control and no super-efficacy attempt)
- Original (Option 1) proposal was single phase 3 design
  - however it was more similar in spirit to seamless Ph2/3 because the part of study prior to first futility analysis was playing the role of Ph2
- It was helpful to separate that part into a separate Ph2b trial with its own objective
  - proof of minimal efficacy in a select sub-population
  - and to size it appropriately (26 events rather than 16)
- More ambitious super-efficacy claim in multiple surgery types would be addressed in a separate Ph3 study

What happened next?

- The proposed GSD approach was generally well accepted
- It prompted the clinical team to formally evaluate operating characteristics of early stopping decisions rather than use some ad-hoc rules
- Even though the proposed Ph2 study was relatively modest in size (~2K vs. ~10K patients), it was still a huge investment
- In such situations teams often are asked to look for ways to accomplish more with that investment

What happened next?

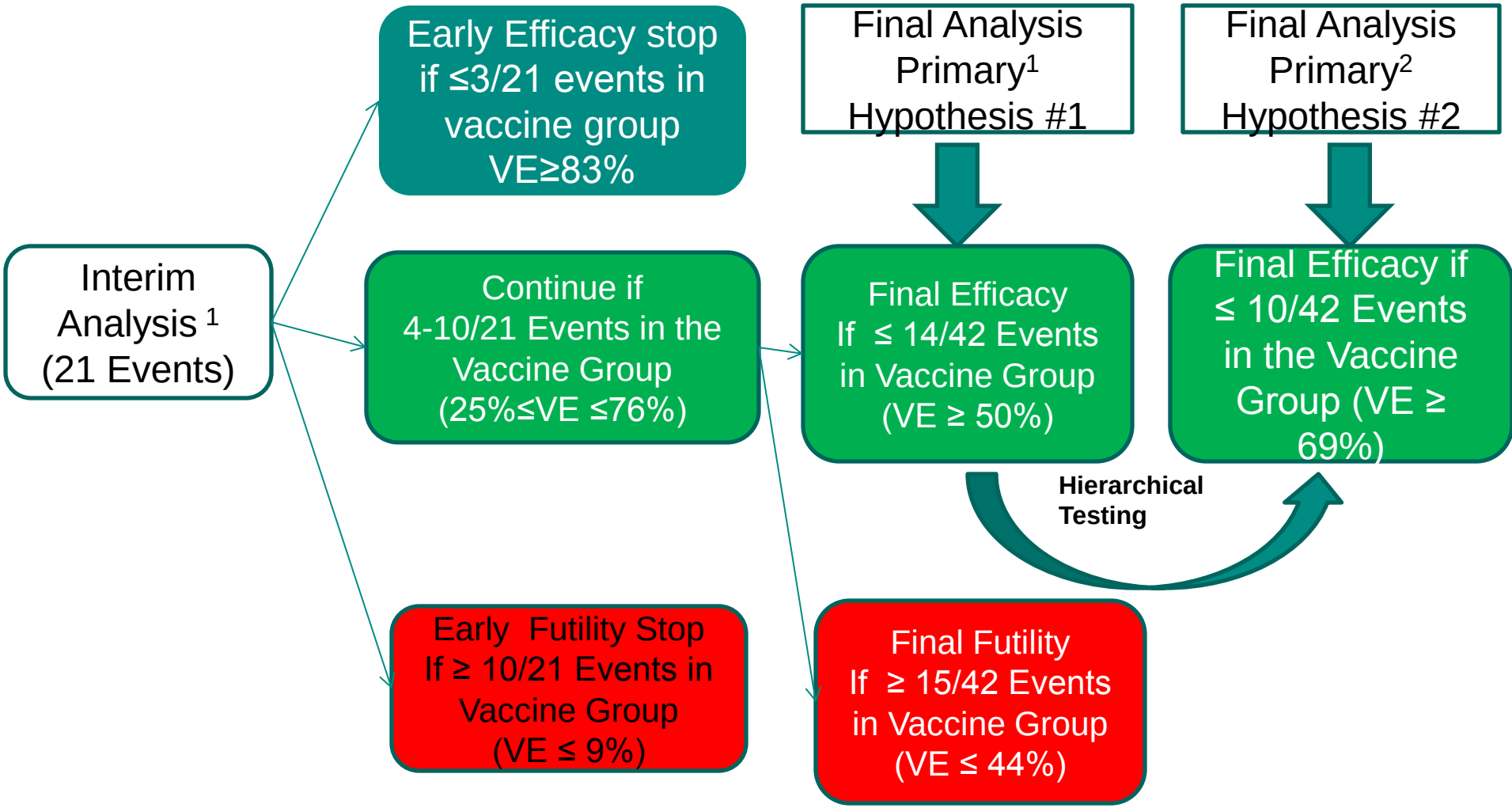
- So the “ph2/3 idea” got re-introduced again but with a different twist:
  - Can we keep the design but up the size it up to make type1/2 error control in-line with Ph3 expectations?
- The discussion was prompted by having an early stopping boundary for efficacy after 13 events (≤2/13 events)
  - If the trial was to stop early with such overwhelming efficacy, the data probably could not stand on its own without another large expensive trial to support it in the filing package
  - the team was asked to make the design “pivotal” quality while keeping other elements the same
  - in case the efficacy is overwhelming, we could attempt to make super-efficacy claim as well
  - So that Ph2b could be considered a pivotal trial if the data is really compelling

What happened next?

- New larger Ph2b study was designed, with α=0.025 and 80% power
- The primary focus if interim look was futility
- As for early efficacy stop, the debate whether to keep it went back and forth a few times
  - Pros: if we stop early, we may end up with insufficient data
  - Con: if we don't put it in, DMC may stop the trial anyway if they see "good" interim results

- It was decided to keep early efficacy stop
- Additional objective was added:
  - if minimal efficacy claim is successful (H_0: VE≤0 is rejected)
  - test for super-efficacy as well (H_0: VE≤30%; with appropriate multiplicity control)

Study Design with Interim Analysis



1. Testing for Minimal Efficacy 0%
2. Testing for Minimal Efficacy 30%

Final Design selected

- A sample size of approximately 2594 will be required to achieve 42 Cases.

  - Enrollment of 2224 subjects would be sufficient to generate 42 Events to achieve 1 sided Lower limit of 97.5% Confidence Interval **>0%** with **80% Power** assuming **60% True Vaccine Efficacy** and **3% Attack rate**.

  - Enrollment of 2594 subjects required to generate 42 Events in order to achieve Lower limit of 1 sided 97.5% Confidence Interval **> 30%** with **90% Power** assuming **80% True Vaccine Efficacy** and **3% Attack rate**.

- One Interim Analysis planned to assess Futility and Efficacy after accrual of 21 Events.

Operating Characteristics of final proposed Ph2b design

	Pr(Early Eff. Stop)	Pr (Early Futility Stop)	Pr(Final Success)	Pr (Final Failure)	Overall Prob of Success	Ave. # of Events
VE=0%	0.001	0.668	0.021	0.311	0.021	28.0
VE=10%	0.002	0.576	0.044	0.379	0.046	29.9
VE=20%	0.004	0.468	0.091	0.438	0.095	32.1
VE=30%	0.008	0.349	0.179	0.464	0.187	34.5
VE=50%	0.046	0.125	0.516	0.313	0.563	38.4
VE=60%	0.109	0.050	0.687	0.154	0.797	38.7
VE=70%	0.251	0.012	0.700	0.036	0.952	36.5
VE=80%	0.527	0.001	0.470	0.002	0.997	30.9

Note:

- Testing H0: VE ≤ 0% vs. Ha: VE > 0%, assuming True VE = 60% and 1 sided alpha level of 0.025
- Decision Criteria (@ Interim Analysis) : Early Futility Stop if ≥ 10/21 Events in Vaccine+Early Eff. if≥3/21
- Decision Criteria (@ Final Analysis) : Final Efficacy if ≤ 14/42 Events in Vaccine
  Final Futility if ≥ 15/42 Events in Vaccine

Discussion

- We presented a story of clinical study design for vaccine efficacy study in the area of rare diseases
- Such studies are often complicated by very low background incidence rate of events in control group
- Because it drives the sample size to be huge, even in case of moderate to large treatment effect
- Another factor contributing to large sample size is super-efficacy requirement
  - It is not enough to show that VE>0%, usually a higher threshold such as VE>30% or 40% has to be met for regulatory approval

Discussion (cont.)

- Two statistical approaches are commonly used to design such studies:
  1. Modeling events as binomial distribution , i.e. 2-sample binomial design
  2. Modeling events as Poisson distribution, then conditioning on total number of events and reducing the problem to one –sample binomial , a.k.a. “event-driven design”
- Event-driven design is more common and was used in our program
- It allows study power to be independent of background event incidence rate
  - Study power is a function of VE and total number of events observed
- But it makes study sample size a random variable with expected value depending on the background incidence rate
  - Which makes logistics of study planning more complicated (e.g. budget, timing, enrollment)
- In other words, regardless of which modeling approach is used, there is no way around the complexities arising from low background event rate

Discussion (cont.)

- In such large trials, even small deviations from assumptions made at design stage can lead to costly consequences
  - actual background incidence rate lower than assumed
  - VE lower than assumed
- Clinical development program needs to properly account for uncertainly about parameters of interest rather just focusing on 2 values: null and alternative
  - The uncertainly about VE and λ_0 is only one part of the design problem
  - Other challenges included heterogeneity of population arising from multiple regions and multiple surgery types

Summary

- To tackle these challenges we went through 3 clinical development options:
  1. (Extreme) Very complex “Ph2/3-like” design attempting to address sub-population issue and uncertainty about parameters all at once via adaptation
  2. (Extreme) Very simple and small Ph2b focusing on proof of efficacy only and dropping sub-population and super-efficacy issues . Ph3 is a separate study

Selected

  3. (Middle-ground) Larger Ph2b GS trial keeping other option #2 elements the same;
     - Possibly a “pivotal trial” in case the data is really compelling
     - Separate Ph3 may be GS or include other adaptive elements such as SSR
- The moral of the story is: if you have multiple competing goals and high uncertainty about design parameters, adaptive design is unlikely to solve these problems at once => need to break into steps

Summary (cont.)

- This program was ripe for adaptive design application because there was pressure to accelerate development facing many challenges at the same time
- What we sometimes forget is that adaptive designs are tailored to specific objectives, at times quite narrow; they are not universal " one-size-fits-all" tools
- This experience demonstrated that teasing out separate trial objectives and assigning priorities was helpful in designing a complex clinical program
- This way the operating characteristics could be evaluated more thoroughly and the decisions were more transparent
- In clinical program option #3 selected, there still may be room for adaptation (ph3 work in progress)
- But it will be done in a more thorough manner (unlike Option #1) and after proper Ph2b (giving more definitive read of VE an d λ_0) is in place

Thank you!

References

1. Jennison, C., Turnbull, B. W. (2000). *Group Sequential Methods with Applications to Clinical Trials*. Boca Raton, London, New York, Washington, D.C: Chapman & Hall.
2. J. Nauta (2010) *Statistics in Clinical Vaccine Trials*. Springer –Verlag Berlin Heidelberg