SARS-COV-2 Vaccine Trials
What do we know and when do we know it?

Kert Viele (@KertViele)
NISS MERCK Webinar
January 13, 2021
Why are we here?

- Four large ongoing vaccine trials
  - AstraZeneca, Janssen, Moderna, Pfizer

- All companies have made their protocols publicly available (this is great!)

- We all want a safe and effective vaccine as quickly as possible.
Outline

• What is their general trial design?
  – i.e. “what’s with these interim analyses?”

• What protections do these trials provide?
  • chance of false discovery
  • safety information

• Now that we know what happened (at least for a couple of them)...what did we learn?
Four trials by the numbers...

<table>
<thead>
<tr>
<th>Sponsor</th>
<th>AstraZeneca</th>
<th>Janssen</th>
<th>Moderna</th>
<th>Pfizer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall Sample Size</td>
<td>30,000</td>
<td>60,000</td>
<td>30,000</td>
<td>44,000</td>
</tr>
<tr>
<td>Randomization Ratio (Trmt : Ctrl)</td>
<td>2:1</td>
<td>1:1</td>
<td>1:1</td>
<td>1:1</td>
</tr>
<tr>
<td>Show superiority to VE =</td>
<td>30%</td>
<td>30%</td>
<td>30%</td>
<td>30%</td>
</tr>
<tr>
<td>Maximal events</td>
<td>150</td>
<td>154 (?)</td>
<td>151</td>
<td>164</td>
</tr>
<tr>
<td>Powered for VE =</td>
<td>50%</td>
<td>60%</td>
<td>60%</td>
<td>60%</td>
</tr>
<tr>
<td>Statistical Method</td>
<td>Poisson regression (age as covariate)</td>
<td>Binomial on Event Split</td>
<td>Cox Proportional Hazards</td>
<td>Binomial on Event Split</td>
</tr>
</tbody>
</table>
Sample size

• Trials really need to see events. To see events, we need a large sample size.

• **October** back of the envelope.....In the USA
  – The current “new daily cases rate” is ~40,000
  – So in 100 days we expect to see 4,000,000 cases
  – About 1.25% of the population (320,000,000)

• With 30,000 patients
  – 15,000 per arm
  – 1.25% of 15,000 placebo is 187.5 events
Safety (Shorter Term)

• Short/mid term reactions to injections
  – samples size measures in terms of patients in trial, should be in the thousands even if events accrue rapidly

• These events should be well estimated
Safety (Long Term)

• A declaration of efficacy
  – does not stop the trial
  – patients are still followed up long term (2 yrs)

• However, at time of EUA, long term followup will be limited

• 2 months median followup
  – would prefer this said “for X participants”
Safety (Disease Severity)

• With 150-165 events, all trials monitor for disease severity
  – some tie this to success

• However, it is very hard to accurately estimate disease severity with these sample sizes
Interim Analyses

• All four trials have interim analyses

• At certain points, efficacy can be declared prior to the maximal event count
  – AstraZeneca, Moderna, Pfizer base these analyses on event counts
  – Janssen’s are “at least weekly”, but require certain conditions to start (more on this later…)

Berry Consultants Vaccine Webinar
Interim Analyses

• All four plans protect against false claims

• If any vaccine is truly only 30% effective....
  – each trial has a 2.5% total chance of falsely declaring efficacy (total across all interims)

• Trials all obtain 85-90% power
  – limited chance of missing an effective vaccine
## Interim Analysis Plans

<table>
<thead>
<tr>
<th></th>
<th>AstraZeneca</th>
<th>Janssen</th>
<th>Moderna</th>
<th>Pfizer</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Maximal Events</strong></td>
<td>150</td>
<td>154 (?)</td>
<td>151</td>
<td>164</td>
</tr>
<tr>
<td><strong>Interim Events</strong></td>
<td>75</td>
<td>see below</td>
<td>53,106</td>
<td>32,62,92,120</td>
</tr>
<tr>
<td><strong>Alpha-spending (one-sided)</strong></td>
<td>p&lt;0.00155, p&lt;0.0245 at final</td>
<td>SPRT* Weekly interims</td>
<td>p&lt;0.0002, p&lt;0.0073, p&lt;0.0227 at final</td>
<td>Pr(VE&gt;30%)&gt;0.995 at interims Pr(VE&gt;30%)&gt;0.986 at final</td>
</tr>
<tr>
<td><strong>Overall Type 1 Error Rate</strong></td>
<td>2.5%</td>
<td>2.5%</td>
<td>2.5%</td>
<td>2.5%</td>
</tr>
</tbody>
</table>

Janssen has conditions on when interims start:
- 50% of patients with 2 months followup
- 20 COVID events (at least 5 severe, at least 6 >= 60 yrs old)

After that interims are weekly

If this occurs after 154 events, there is a single analysis when these conditions are met.
Lots of convincing interim data

Black = samples paths from 30% VE
Green = sample paths from 60% VE
While earlier time points present the possibility of black/green overlap, there are lots of results which are firmly good or firmly bad.

Data in these areas are firmly below VE 30%
If VE=80%

A hugely effective vaccine would differentiate itself from the black VE=30% range very quickly.
Pfizer

Table 5. Interim Analysis Plan and Boundaries for Efficacy and Futility

<table>
<thead>
<tr>
<th>Analysis</th>
<th>Number of Cases</th>
<th>Success Criteria(^a)</th>
<th>Futility Boundary</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>VE Point Estimate (Case Split)</td>
<td>VE Point Estimate (Case Split)</td>
</tr>
<tr>
<td>IA1</td>
<td>32</td>
<td>76.9% (6:26)</td>
<td>11.8% (15:17)</td>
</tr>
<tr>
<td>IA2</td>
<td>62</td>
<td>68.1% (15:47)</td>
<td>27.8% (26:36)</td>
</tr>
<tr>
<td>IA3</td>
<td>92</td>
<td>62.7% (25:67)</td>
<td>38.6% (35:57)</td>
</tr>
<tr>
<td>IA4</td>
<td>120</td>
<td>58.8% (35:85)</td>
<td>N/A</td>
</tr>
<tr>
<td>Final</td>
<td>164</td>
<td>52.3% (53:111)</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: IA = interim analysis; N/A = not applicable; VE = vaccine efficacy.
Note: Case split = vaccine : placebo.
a. Interim efficacy claim: P(VE >30%|data) > 0.995; success at the final analysis: P(VE >30%|data) > 0.986.
Pfizer Efficacy Rules

Pfizer will stop and declare efficacy in the red rectangles

All are below the vast majority of black (VE=30%) paths. They have a collective 2.5% type I error
What is more compelling?
6-26 or 53-111

6-26 stochastically dominates 53-111

6-26 has more variability
(we “know” >30% but 60% and 90% are very different)

Should these questions be “punted” in a public health emergency?
What about safety?

• Stopping at 32 (or 53, 62, or 75…or) does limit the amount of available safety data.

• For short term injection reactions
  – sample size is the number of patients, not events
  – if sponsor had enrolled thousands, even with 32 events this sample size could be quite large.

• For long term followup, the real question is the timing of the stop.
Long term followup

• Important to have sufficient followup

• Unclear what is “sufficient”

• Some discussion about having 50% or 100% of the patients followed up 2 months.

• Note diminishing returns of increasing N.
  – Suppose an AE rate was estimated 8% +/- 3.5%
  – Double the sample size, that becomes 8% +/- 2.5%
What happened?

• Pfizer bypassed first interim (FDA request for more extensive safety data)
• Upon agreement with FDA, Pfizer ran the second interim, but ended up with more than 62 events
  – there should be a plan somewhere
• VE \sim 90\%, should be no controversy VE>30\%
Predicting N=164 from interim

Was this right? It actually appears pessimistic….next result with over 164 events had 95% efficacy, perhaps indicating increasing immunity over time.