Mitigating “force majeure” trial disruptions by leveraging external data

With special reference to pooling approaches

Marc Vandemeulebroecke, NISS Ingram Olkin Forum Series, 23 March 2021
Credit to **Eva Hua**, Frank Bretz, Paul Gallo, Reinhold Janocha, Thomas Severin, Jiawei Wei, Dong Xi

NOVARTIS | Reimagining Medicine

Force majeure trial disruption

STATISTICS IN BIOPHARMACEUTICAL RESEARCH
2020, VOL. 12, NO. 4, 414–418
https://doi.org/10.1080/19466315.2020.1785931

Taylor & Francis
Taylor & Francis Group

Under a Black Cloud Glimpsing a Silver Lining: Comment on Statistical Issues and Recommendations for Clinical Trials Conducted During the COVID-19 Pandemic

Rob Hemmings

Supportive approaches might look to integrate data from external sources, to augment the control arm, or pool trial data with results from previous studies but regulators will have to consider approvals based on a greater than usual degree of uncertainty and use relevant and feasible post-authorization data generation to complement the pre-authorization study or studies. In the

STATISTICS IN BIOPHARMACEUTICAL RESEARCH
2020, VOL. 12, NO. 4, 419–426
https://doi.org/10.1080/19466315.2020.1788984

Taylor & Francis
Taylor & Francis Group

Challenges in Assessing the Impact of the COVID-19 Pandemic on the Integrity and Interpretability of Clinical Trials

Mouna Akacha[a], Janice Branson[a], Frank Bretz[a,b], Bharani Dharan[c], Paul Gallo[c], Insa Gathmann[a], Robert Hemmings[d], Julie Jones[a], Dong Xi[c], and Emmanuel Zuber[a]

Other approaches to compensate for lost information can be considered as well, such as leveraging data on short-term endpoint(s) that are correlated with the primary response or to integrate data from external sources, to augment the control arm, or pool trial data (Hemmings 2020).

STATISTICS IN BIOPHARMACEUTICAL RESEARCH
2020, VOL. 12, NO. 4, 399–411
https://doi.org/10.1080/19466315.2020.1779122

Taylor & Francis
Taylor & Francis Group

Statistical Issues and Recommendations for Clinical Trials Conducted During the COVID-19 Pandemic

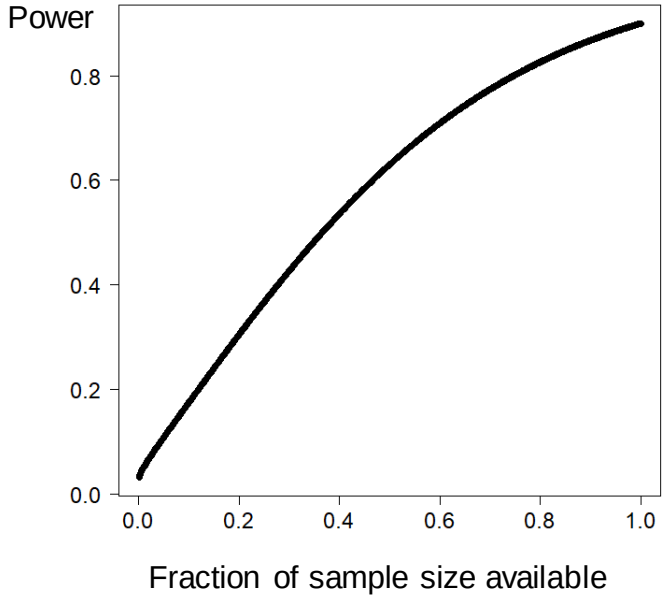
R. Daniel Meyer[a], Bohdana Ratitch[b], Marcel Wolbers[c], Olga Marchenko[d], Hui Quan[e], Daniel Li[f], Christine Fletcher[g], Xin Li[h], David Wright[i], Yue Shentu[j], Stefan Englert[k], Wei Shen[l], Jyotirmoy Dey[m], Thomas Liu[n], Ming Zhou[f], Norman Bohidar[o], Peng-Liang Zhao[e], and Michael Hale[p]

For some trials, it may not be feasible to increase sample size and the trial will fall short of enrollment target. Given the extraordinary circumstances, we advocate more flexibility to consider methods for quantifying evidence across multiple trials and sources, including use of historical control arm data and real-world data, although sources and methodology for selection of such data would need to be planned and agreed with regulatory agencies in advance. If the observed treatment

How much power do I lose?

- Level α test with power $1-\beta$
- Normally distributed test statistic
- Fraction p of targeted information available

  ➔ $\text{Power} = P\left(Z > \left(1-\sqrt{p}\right) z_\alpha - \sqrt{p}\, z_\beta\right)$
- Example with $\alpha = 0.025$ and $1-\beta = 90\%$ ➔



NOVARTIS | Reimagining Medicine

Kunz et al. (2020), Akacha et al. (2020)

How could I compensate for it?

- Leveraging short-term endpoint data ➔ Previous talk
- Leveraging external information ➔ This & next talk
- More ➔ ...?

What external data sources are there?

Historical

- Completed studies from the same development program (Phase II...)
- From other programs / sponsors
  - Typically on control / placebo
  - Often more data
  - Sometimes older
- Real-world data

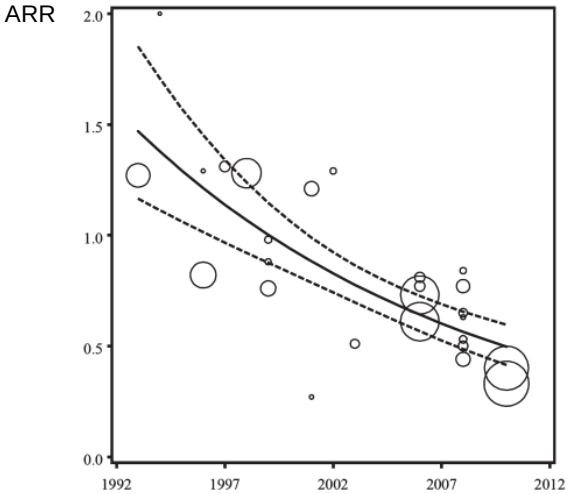
Concurrent

- Parallel studies from the same development program
  - Replicate Phase III
  - Head-to-head
- Real-world data

Which points should I consider?

- Time trends
- Differences in
  - Population
  - Endpoints
  - Treatment regimens
  - Regions
  - Co-medications
  - Other protocol aspects

Annualized Relapse Rate (ARR) on Placebo in Relapsing-Remitting Multiple Sclerosis



Nicholas et al. (2011)

How can I bring the data together?

- Pooling 

This talk

- Regression / adjustment for covariates
- Meta-analysis & accommodating unidentified /-able variation 
  - In particular, Meta-Analytic Predictive (MAP) approaches; Schmidli et al. (2020)
- Propensity score methods
  - E.g., next talk; Schmidli et al. (2020)
- Modeling & Simulation, incl. PBPK, QSP...
  - E.g., Geerts and Van der Graaf (2020)

MAP approach

- Typically used for leveraging historical (not concurrent) information to complement the current trial
- Accounts for between-trial variability by discounting information, rarely by modeling covariates
- Has been accepted by Health Authorities in certain situations
- Method
  - Historical trials with true mean effects $\theta_1, \ldots, \theta_K$
  - Effect in new study: θ^*
  - Model: $\theta_1, \ldots, \theta_K, \theta^* \sim N(\mu, \tau^2)$

Neuenschwander et al. (2010)

MAP approach – examples

- Proof of Concept in Crohn's disease

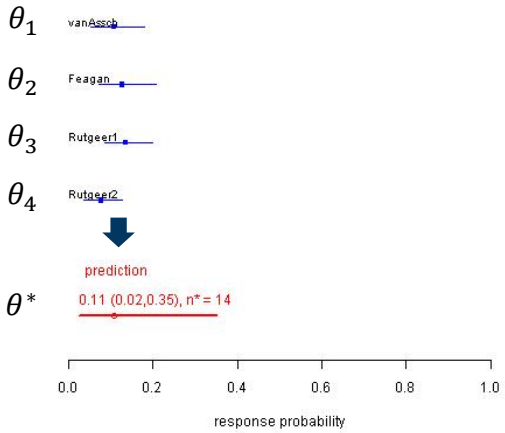
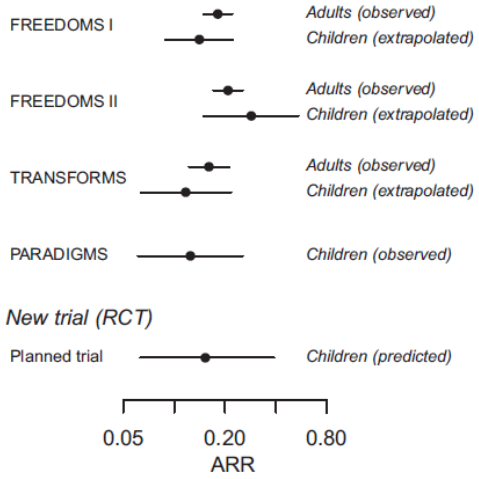


Table 2 Primary end point (Bayesian analysis including historical controls)

CDAI change from baseline to week 6			
Treatment	**n**	**Mean (SD)**	**95% Credible interval***
Secukinumab twice 10 mg/kg	39	−29.2 (14.0)	−56.9 to −1.4
Placebo	20	−63.1 (13.9)	−90.4 to −35.9
ΔCDAI (AIN457 vs placebo)		33.9 (19.7)	−4.9 to 72.9

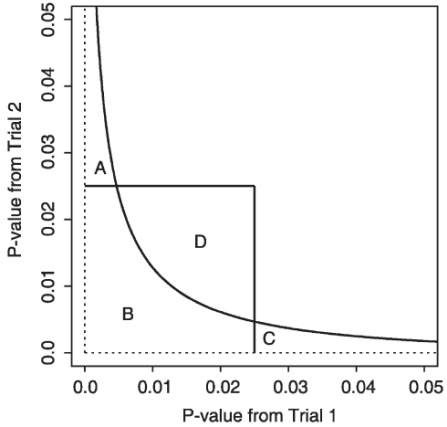
- Pediatric study in Multiple Sclerosis
  - In FDA's Complex Innovative Designs program



Hueber et al. (2012), Schmidli et al. (2020)

Pooling

- Pooling approaches have been discussed even without trial disruptions
  - Two pivotal trials, each at $\alpha = 0.025$
  - One large trial at $\alpha = 0.000625$ provides the same false positive protection
  - Both trials powered at 90% ➔ Program power* = $0.9^2 = 81\%$
  - But a pooled analysis at $\alpha = 0.000625$ has 91% power!



Pooled		Two trials
P(A,B,C)	$>_{H1}$	P(B,D)

*Probability that both trials succeed under the design alternative

Maca et al. (2002)

Pooling

- Pooling approaches have been discussed even without trial disruptions
  - Two pivotal trials, each at $\alpha = 0.025$
  - One large trial at $\alpha = 0.000625$ provides the same false positive protection
  - Both trials powered at 90% ➔ Program power* = $0.9^2 = 81\%$
  - But a pooled analysis at $\alpha = 0.000625$ has 91% power!
- Can we leverage this in case of trial disruption?
  - Say, the trials are curtailed at 85% of the planned overall sample size
  - Pooled analysis at $\alpha = 0.000625$ has power 84%
  - Additionally, $\alpha = 0.000625$ could be relaxed, e.g. for secondary objectives
- Prerequisite: «Poolability», i.e. same or very similar protocol (population, endpoints, treatment regimens etc.; see above)

NOVARTIS | Reimagining Medicine

*Probability that both trials succeed under the design alternative

Maca et al. (2002)

Case study

- Two identical pivotal RCTs

- Multiple hypotheses tested in each, e.g. «H1hp»
- Sufficiently powered for primary endpoint, high & low dose vs. active control

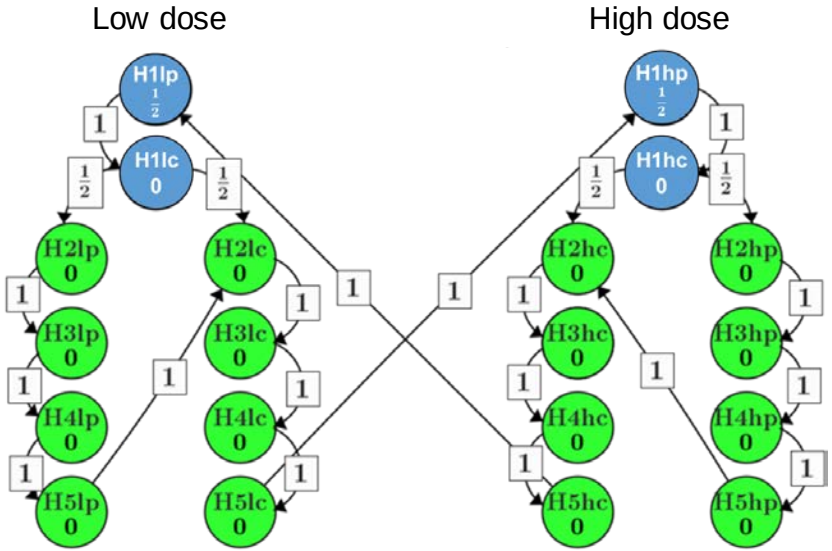
1: primary, **2-5**: secondary
h: high dose arm, **l**: low dose arm
p: vs. placebo, **c**: vs. active control

NOVARTIS | Reimagining Medicine

Hua et al. (2021)

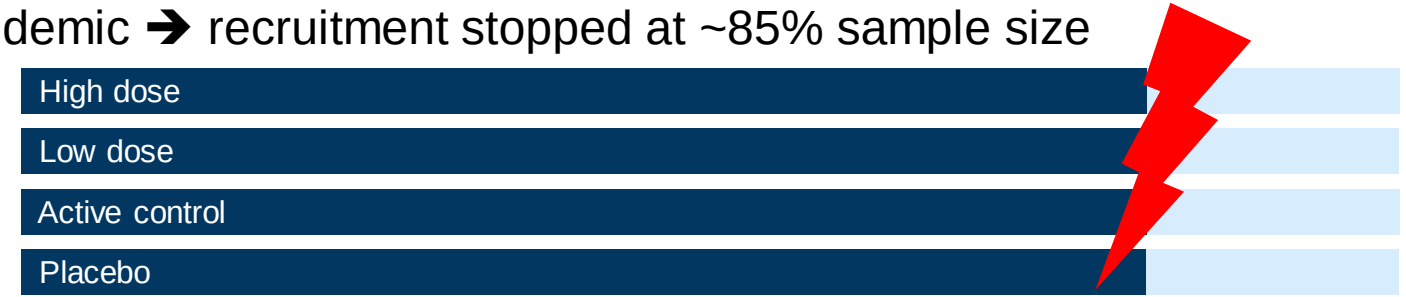
Case study

- Original statistical testing strategy (same in each study)



Case study

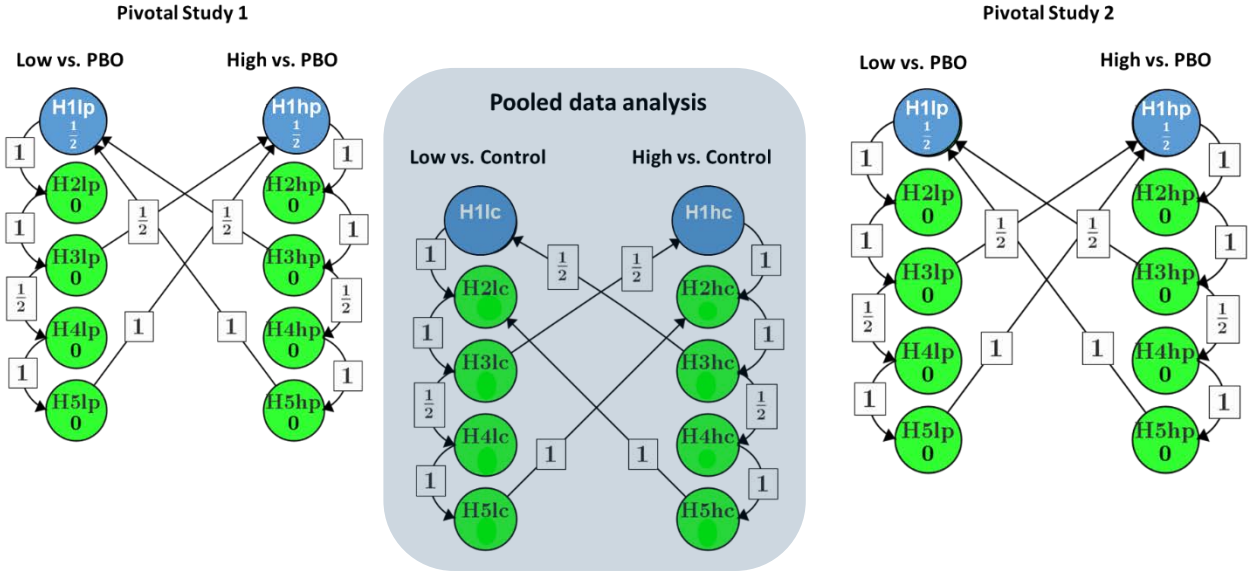
- Pandemic ➔ recruitment stopped at ~85% sample size



- Placebo comparisons remain sufficiently powered
- But not the comparisons vs. active control!

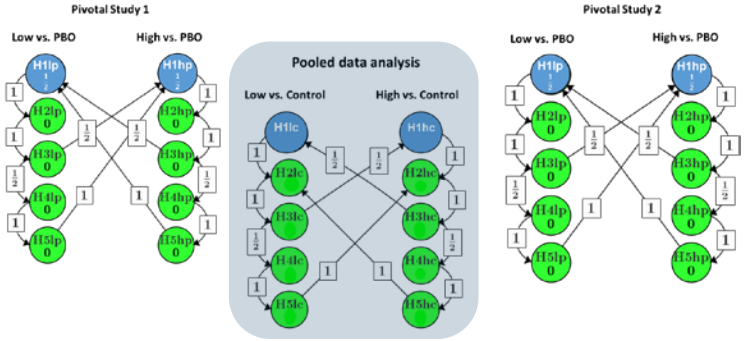
Case study

- (Initially) proposed mitigation: **Perform tests vs. active control in pool**



Case study

- (Initially) proposed mitigation: **Perform tests vs. active control in pool**



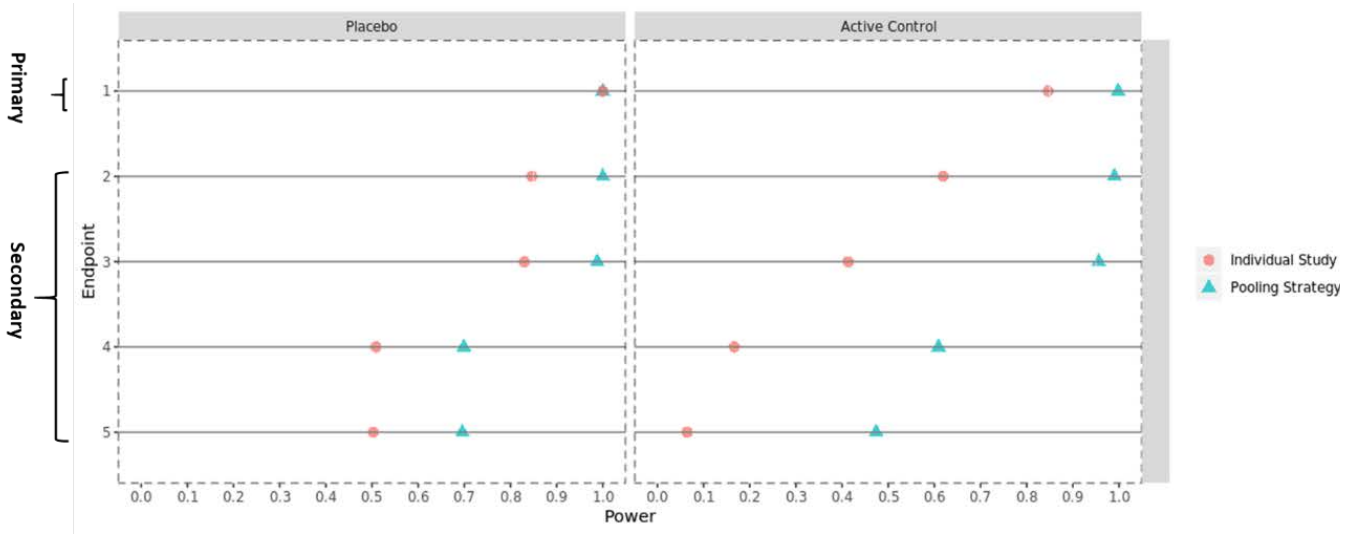
- Additional requirement: Any endpoint can only be tested vs. active control in the pool if it shows superiority in each of the separate studies vs. placebo
- Initial type I error for tests vs. active control: $\frac{\alpha}{2} - \alpha^2 = \frac{0.025}{2} - 0.025^2$ per dose

  ➔ Type I error controlled at <0.025 at study level for all placebo comparisons, at <0.025 at submission level for all hypotheses

Bretz et al. (2019)

Case study

- Program power*: **Great gains for secondary endpoints vs. active control**



NOVARTIS | Reimagining Medicine

*Power at the submission level with ~85% of the planned sample size

Case study – Discussion

- It's not just a numbers game
  - **Replication** is important – and can still be provided by additional analyses per study
  - **A-priori specification** is a must
  - **Transparency** (also towards Health Authorities) is crucial
- There's more to the story
  - Logistics, estimands / intercurrent events, sensitivity analyses, etc...
- And it continues
  - Recruitment restarted during the pandemic and could be completed
  - Yet, pre- vs. during-pandemic heterogeneity cannot be excluded
  - Final proposal is different from the one presented here

Conclusions

- **Leveraging external data** may help mitigate the impact of «force majeure» trial disruptions
- Consider **various types of data sources** (historical / concurrent, clinical trials / real-world, etc...)
- **Different quantitative approaches** (pooling, regression, MAP, propensity scores, M&S) are available depending on the situation
- **Be aware** (and beware) of time trends and differences in population, endpoints, treatment regimens, etc...
- **Extraordinary circumstances may require extraordinary measures**
  - While still providing «**substantial evidence of effectiveness**» (FDA, 1998) – and sufficient **safety** information as well!

References (1)

- Akacha et al.: Challenges in Assessing the Impact of the COVID-19 Pandemic on the Integrity and Interpretability of Clinical Trials. Stat Biopharm Res (2020).
- Bretz et al.: A General Graphical Approach to Sequentially Rejective Multiple Test Procedures. SiM (2009).
- Bretz et al.: Replicability, Reproducibility, and Multiplicity in Drug Development. Chance (2019).
- FDA: Guideline for Industry: Providing Clinical Evidence of Effectiveness for Human Drug and Biological Products. FDA (1998)
- Geerts and Van der Graaf: A Modeling Informed Quantitative Approach to Salvage Clinical Trials Interrupted due to COVID-19. Alzheimer's and Dementia (2020).
- Hua et al.: A Phase 3 Trial Analysis Proposal for Mitigating the Impact of the COVID-19 Pandemic. To appear in Stat Biopharm Res (2021).
- Hemmings: Under a Black Cloud Glimpsing a Silver Lining: Comment on Statistical Issues and Recommendations for Clinical Trials Conducted During the COVID-19 Pandemic. Stat Biopharm Res (2020).
- Hueber et al.: Secukinumab, a human anti-IL-17A monoclonal antibody, for moderate to severe Crohn's disease: unexpected results of a randomised, double-blind placebo-controlled trial. Gut (2012).

References (2)

- Kunz et al.: Clinical Trials Impacted by the COVID-19 Pandemic: Adaptive Designs to the Rescue? Stat Biopharm Res (2020).
- Maca et al.: Reconsidering some Aspects of the Two-Trials Paradigm. J Biopharm Stat (2002).
- Meyer et al.: Statistical Issues and Recommendations for Clinical Trials Conducted During the COVID-19 Pandemic. Stat Biopharm Res (2020).
- Neuenschwander et al.: Summarizing Historical Information on Controls in Clinical Trials. Clin Trials (2010).
- Nicholas et al.: Trends in annualized relapse rates in relapsing–remitting multiple sclerosis and consequences for clinical trial design. MSJ (2011).
- Schmidli et al.: Beyond Randomized Clinical Trials: Use of External Controls. CPT (2020).