Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
COVID-19 hits a trial: Arguments against hastily deviating from the plan

Anika Großhennig*, Armin Koch*

Institut für Biometrie, Medizinische Hochschule Hannover, Germany

ABSTRACT

The COVID-19 pandemic has substantially impacted the conduct of clinical trials. While initially preparing for a period of time, where it would likely be impossible to supervise trials in the usual way and precautionary measures had to be implemented to care for medication supply and general safety of study participants it is now important to consider, how the impact of the pandemic on trial outcome can be assessed, which measures are needed to decide, how to proceed with the trial and what is needed to compensate for irregularity introduced by the pandemic situation. Obviously not all trials will suffer to the same degree: some trials may be close to finalizing recruitment, others may not yet have started. Similarly not all clinical trials investigate vulnerable patient populations, but some will and may in addition have recruited to an extent that beneficial effects achieved in the initial phase of the trial may be outweighed by an increase e.g. in mortality that impacts both treatment groups. The situation is further complicated by the fact that the pandemic reached different countries in the world and even cities in one country at different points in time with different severity. Our example is a randomized and double-blind clinical trial comparing digitoxin and placebo in patients with advanced chronic heart failure. This trial has recruited roughly 1/3 of the overall 2200 patients when the disease outbreak reached Germany. We discuss how simulations and theoretical considerations can be used to address questions about the need to increase the overall sample-size to be recruited to compensate for a potential shrinkage of the treatment effect caused by the COVID-19 pandemic and what role the degree of consistency could play when comparing pre-, during- and post- COVID-19 periods of trial conduct regarding the question, whether the treatment effect can be considered consistent and with this generalizable. This is dependent on the size of the treatment effect and the impact of the pandemic. We argue, that in case of doubt, it may be wise to proceed with the original study plan.

1. Introduction

Currently the COVID-19 pandemic challenges all aspects of our life as individuals, as citizens of a country and part of a world, having had benefit from co-operation and personal exchange for a long time. Long incubation times, limited symptomatology in many, easy spread of the infection, limited ability to prevent infection in the individual beyond social distancing and last, but not least, the lack of effective prevention and treatment lead to the sequential outbreak of disease in all regions of the world.

There is also no precedence for an event with so direct impact on worldwide clinical trials research as the current COVID-19 pandemic. At best, some regions had the opportunity to prepare for a time-span where patients cannot attend the pre-planned trial visits and limited or no access of study patients to the healthcare system in charge of taking care of their well-being. In the field of drug trials strategies for medication supply had to be developed and concepts were needed, how to obtain a limited set of information to assure patients’ safety in the first place and consider secondly, how to document efficacy of treatment with extremely reduced remotely achieved data. In some instances, the impact of the pandemic on the trial can only be followed by assessing the amount of missing values in the structured trial documentation.

After this first step now aspects of analysis and interpretation become more important. It is anticipated that only few trials will be so close to the start of recruitment that they simply can be put on hold and re-opened after the pandemic. Likewise few trials are so close to completion that they can be prematurely stopped with good expectation that this will not negatively impact on the interpretation of trial outcome. These two categories will not be a major problem for the discussion. The discussion will have to center around those trials that are at a state, where neither a re-start, nor early termination are reasonable options.

The situation is even more complicated as COVID-19 doesn’t hit the trial in all regions and centers at the same time. However, start (and
hopefulness ending) of specific measures can be identified. Trial data need to be presented in a way that allows assessing the impact of the pandemic in a structured and informative way. The impact of the COVID-19 pandemic on decision making regarding the further conduct of the trial requires careful consideration: a simple stop for futility wastes all previous investments into trial setup. More importantly, premature termination would mean that efforts of patients willing to contribute to generate in a human experiment knowledge regarding their disease for future patients, would have been in vain.

The decision for a potential resizing of a trial because of an anticipated higher variability in outcome, or for continuation of the trial as pre-planned, require careful planning and consideration to be properly implemented into the design. Careless activity may damage trial integrity and create obstacles to interpretation that are difficult to resolve. For this reason regulatory guidance emphasizes that in first place standard monitoring strategies for blinded data are used to assess the impact of the pandemic and, where access to unblinded data is felt to be important for decision making, this should preferably be done by an independent body (e.g. a well-trained DMC). For obvious reasons such decision making is far more complicated for open label clinical trials or single-arm clinical trials.

EMA and FDA came up very fast with regulatory guidance and points to consider-documents discussing the implications on methodology aspects of ongoing clinical trials [1,2]. These documents recommend the conduct of additional statistical analyses without being specific about their type and nature. A discussion on specific scenarios that can occur and how to handle them is also not included. In this paper we want to suggest some assessment strategies that may be helpful to assess the ability of the trial to provide robust conclusions about the therapeutic question at hand despite the impact of the pandemic. We will discuss such considerations based on a randomized and double-blind clinical trial comparing Digitoxin to Placebo in patients with advanced chronic heart failure (DIGIT-HF).

2. Simulation study and results

The motivating example for our simulation study is the DIGIT-HF study which is a large multicenter, randomized, double-blind, placebo-controlled investigator initiated clinical trial. The full description of the design of the DIGIT-HF trial is depicted and discussed elsewhere [3]. The primary objective of this trial is to demonstrate that digitoxin is effective in patients with heart failure.

The primary endpoint of the study is time to first hospitalization due to worsening heart failure or death, whichever occurs first. At the end of February 2020 roughly 1/3 of the overall 2200 patients were randomized.

For the sake of convenience we assume that studies have a history before, during and after COVID-19 pandemic. In reality this is even more complicated because patients may have a lower or higher risk to be infected and to suffer from severe consequences of an infection, and the influence of the COVID-19 pandemic occurs at different study sites to a different extend and at different time points. Risks may be posed to the study patients due to a reduced medical oversight and limited access to study centers during a lock down. Additional adjustments are needed to clearly describe the effect of the pandemic beyond the usual time-axis counting study time since randomization of individual patients.

A realistic simulation model will have to incorporate all these aspects for a survival study with a time to event endpoint as DIGIT-HF. The constant linear hazard model which is usually assumed for survival studies has to be replaced by piecewise (constant) hazard functions that cover the effects due to different study phases: pre, during and post the COVID-19 pandemic.

Beside this piecewise constant hazard function modelled with different assumptions for the exponential distribution in different study phases, for each patient the starting- and endpoint for the pandemic may be modelled with, for example a random offset generated with a uniformly distributed random number. The concrete technical implementation will also have to define parameters for this complex simulation model which is a challenging exercise on its own. Therefore we decided for a rather crude simulation model which allows us at the one hand to approximate the potential problems and which is at the other hand suitable to estimate the maximal impact of the pandemic.

We investigated the baseline risk rates in the three trial periods and ignore the time under observation of patients randomized before, during, or after the pandemic. We modified the baseline event rate during the pandemic and the treatment effect, but did not further investigate, whether the different components of the composite endpoint death or re-hospitalization for worsening heart failure would be affected differently: there were considerations that due to the pandemic patients would not visit the hospital even in the event of worsening heart failure [4], but we felt that this would overall challenge the importance and clinical relevance of this component of the composite endpoint.

Simulations were performed with SAS software (Version 9.4, SAS Institute Inc., Cary, NC, USA).

For each time period (pre, during, post pandemic) events of the primary endpoint in the treatment and the control group were generated on the basis of Bernoulli distributed random numbers. In the analysis we performed logistic regression model including treatment group and study period. For the assessment we calculated the empirical power defined as the number of simulations where the respective two-sided p-value of the overall treatment effect in the logistic regression model was smaller than 5% suggesting superiority of the experimental treatment over control in a superiority study.

Assumptions for the baseline event rates were based on the initial sample size calculation of the DIGIT-HF study which had been based on the results of the SHIFT study observing 26% vs 31% patients suffering a primary endpoint event within two years for patients who received digitoxin or placebo [5].

In addition we assumed that the baseline risk after the pandemic is the same as before the pandemic.

We considered 6 scenarios: 1) the null-hypothesis with no treatment effect, 2) no change between the three time periods, 3) increase of the baseline event rates by 8% during COVID-19 which is in line with the reported excess mortality in Germany in April 2020 [6], 4) increase of the event rates by 15% during COVID-19 which is in line with the excess mortality reported from countries that are more affected by COVID-19, 5) decrease in the baseline event rates by 4% to address the possible but not probable scenario that less hospitalizations at the end cause a cumulative reduction of the rate of the primary endpoint [4], and 6) the scenario that the increase in the baseline event rate reduced the treatment effect to zero during the pandemic.

In the simulation scenarios the assumptions for the baseline event rates of the control group, the treatment effect measured as odds ratio, as well as the overall sample size of 2600 patients were made in line with the initial assumptions for the DIGIT-HF trial. For the scenarios 3–5 we examined different distribution of the sample size for the three periods to investigate the impact of different durations of the pandemic. For each scenario a total of 1000 simulation-runs were performed.

The detailed results of the simulations are reported in Table 1.

Overall, the simulated increases in the baseline event rate (mortality or re-hospitalization for worsening heart failure) due to the COVID-19 pandemic has in our simulations no major effect on the overall empirical power of the clinical trial as long as (i) the overall sample-size is kept as pre-planned and (ii) the absolute treatment effect is not affected.

The first two scenarios were included to validate the simulation. As expected for scenario 1, which projects the null-hypothesis with no treatment effect, we observed an empirical power of about 5%. Scenario 2 depicts the study without impact of the pandemic and the empirical power is slightly larger than 80%. Surprisingly the power is quite robust against different sample sizes during pandemic (i.e. the proportion of the trial that is affected) and against increase and decrease of baseline
unblinding a trial in a hurry, or to impatiently decide, how to proceed. We explicitly caution against is most informative for decision making. However, this is not necessarily reasonable to proceed with a trial that has been hibernated during the pandemic. Obviously an unblinded look into the study data during the pandemic. It is obvious that specific indicators need to be discussed and developed that can model the period, where COVID-19 pandemic may have impacted on the patients from this center / region. A close inspection of blinded event rates over time and after adjustment and modelling the COVID-19-period can be informative about immediate intervention effects. If nothing is seen there after careful inspection, the need to unblind the data may not be given. In case of doubt, unblinding may be indeed necessary, but this should be carefully pre-planned as any interim analysis and the respective assessment should be handed over to an independent body like the data safety monitoring board (DSMB). “Do nothing” and proceed as initially planned is a valid option: early termination may increase difficulties in trial assessment because limiting the amount of information that is available for the assessment of benefits and risks. There is a reduction in power if the treatment effect is reduced during the COVID-19 pandemic. Proceeding, however, with the trial into a post-pandemic era, and demonstrating descriptive homogeneity of effects before and after the pandemic may put formal significance of the trial into perspective if this trial provides in all other aspects robust estimates for the treatment effects in relevant subgroups for benefit/risk assessment, as well.

The main recommendations are summarized in box 1. It is obvious that specific indicators need to be discussed and developed that may allow to gauge the impact of the pandemic on a specific trial. We encourage the development of simulation models to investigate the impact in various scenarios. The ability to describe on a center level, where the impact of the pandemic started (and ended) is relevant for decision making. Moving averages, kernel density estimates, sliding window analyses, change-point models can help to address the question, whether the COVID-19 pandemic did impact the patients from this center / region. Additionally trial patients (at least in Germany) are usually particularly well informed and it is therefore plausible that they belong to those, who carefully respect social distancing to self-protect for their safety and the safety of others. In the specific setting under investigation with recurrent worsening of heart failure there may be an initial suspicion that in the pandemic situation patients may not attend the hospital, but if this would have an impact on the estimated treatment effect, this would challenge the clinical relevance of the definition of worsening heart failure and as a component of the primary endpoint.

### What could / should be done?

- Based on center and region information covariates should be developed that can model the period, where COVID-19 pandemic may have impacted on the patients from this center / region.
- A close inspection of blinded event rates over time and after adjustment and modelling the COVID-19-period can be informative about immediate intervention effects. If nothing is seen there after careful inspection, the need to unblind the data may not be given.
- In case of doubt, unblinding may be indeed necessary, but this should be carefully pre-planned as any interim analysis and the respective assessment should be handed over to an independent body like the data safety monitoring board (DSMB).
- “Do nothing” and proceed as initially planned is a valid option: early termination may increase difficulties in trial assessment because limiting the amount of information that is available for the assessment of benefits and risks.
- There is a reduction in power if the treatment effect is reduced during the COVID-19 pandemic. Proceeding, however, with the trial into a post-pandemic era, and demonstrating descriptive homogeneity of effects before and after the pandemic may put formal significance of the trial into perspective if this trial provides in all other aspects robust estimates for the treatment effects in relevant subgroups for benefit/risk assessment, as well.

### 3. Discussion and recommendations

The current situation doesn't have precedence in the history of clinical trials research: studies were terminated because of recruitment problems and turned out to meet their primary objective, other studies were terminated because in the pre-planned setting the likelihood was too low that the trial could meet its objective. Never before, however, external influences, completely unforeseen at the planning stage, questioned the success of trials on a global level. In respect of patient safety, which is always highest priority in clinical trials, visits to study centers, diagnostic measures and other personal contacts to the investigators had to pause and only those measures, important to guarantee the safety of treated patients (medication supply, urgently needed assessments in a distanced form) were possible.

The study in our example meanwhile has restarted recruitment and region by region missed study visits and patient investigations are scheduled. It is the responsibility of Sponsors and investigators to ask themselves, in how far the consequences of the interruption of the trial can be mitigated and whether it is still ethically justified and scientifically reasonable to proceed with a trial that has been hibernated during the pandemic. Obviously an unblinded look into the study data is most informative for decision making. However, this is not necessarily in the interest of the trial integrity. We explicitly caution against unblinding a trial in a hurry, or to impatiently decide, how to proceed. It is an old rule that trials should be left untouched as long as possible.

On the whole we are confident that at least for the randomized clinical trials, the impact of the COVID-19 pandemic is limited. Our simplifying simulations indicate, that as long as COVID-19 pandemic would only impact on the baseline-risk, but would leave the treatment effect unaffected, there is (if at all) limited impact on the power of the trial as long as the recruitment continues to the pre-planned sample size.

We also see that, if for a limited period of time the treatment effect would be zero as a consequence of increases in overall mortality, the impact on the trials ability to meet its objective is likewise relatively small. Additionally trial patients (at least in Germany) are usually particularly well informed and it is therefore plausible that they belong to those, who carefully respect social distancing to self-protect for their safety and the safety of others. In the specific setting under investigation with recurrent worsening of heart failure there may be an initial suspicion that in the pandemic situation patients may not attend the hospital, but if this would have an impact on the estimated treatment effect, this would challenge the clinical relevance of the definition of worsening heart failure and as a component of the primary endpoint.

### What could / should be done?

- Based on center and region information covariates should be developed that can model the period, where COVID-19 pandemic may have impacted on the patients from this center / region.
- A close inspection of blinded event rates over time and after adjustment and modelling the COVID-19-period can be informative about immediate intervention effects. If nothing is seen there after careful inspection, the need to unblind the data may not be given.
- In case of doubt, unblinding may be indeed necessary, but this should be carefully pre-planned as any interim analysis and the respective assessment should be handed over to an independent body like the data safety monitoring board (DSMB).
- “Do nothing” and proceed as initially planned is a valid option: early termination may increase difficulties in trial assessment because limiting the amount of information that is available for the assessment of benefits and risks.
- There is a reduction in power if the treatment effect is reduced during the COVID-19 pandemic. Proceeding, however, with the trial into a post-pandemic era, and demonstrating descriptive homogeneity of effects before and after the pandemic may put formal significance of the trial into perspective if this trial provides in all other aspects robust estimates for the treatment effects in relevant subgroups for benefit/risk assessment, as well.

### Simulation parameters for the pre pandemic period are based on the assumptions for the initial sample size calculation: Baseline risk rate of the control group: 31%, Treatment effect measured as odds ratio: 1.0 for scenario 1 / 0.78 for all other scenarios, sample size: 800.

Simulations parameters for the post pandemic period are the same as in the pre pandemic period, except for sample size: sample size was set in dependence to the sample size that was chosen in the during pandemic period to achieve an overall sample size of 2600 patients (a-scenario: 1400, b-scenarios: 1200, c- scenarios: 600).

### Risk rates (scenarios 3, 4, 5).

| Scenario | Simulation parameters during pandemic | Empirical power (%) |
|-----------|----------------------------------------|---------------------|
|           | Sample size | Baseline risk rate | Odds ratio | Pre | During | Post | Total |
| 1-A       | 400         | 0.31               | 1.00       | 5.1 | 5.5    | 5.4  | 6.0   |
| 2-A       | 400         | 0.31               | 0.78       | 36.9| 18.6   | 38.3 | 59.2  |
| 3-A       | 400         | 0.34               | 0.78       | 34.4| 20.4   | 58.1 | 82.0  |
| 3-B       | 600         | 0.34               | 0.78       | 33.3| 33.2   | 48.5 | 83.0  |
| 4-A       | 1200        | 0.34               | 0.78       | 35.2| 35.3   | 26.3 | 83.0  |
| 4-B       | 400         | 0.41               | 0.78       | 35.8| 24.2   | 55.6 | 84.0  |
| 5-A       | 400         | 0.41               | 0.78       | 33.5| 35.7   | 52.2 | 84.5  |
| 5-B       | 600         | 0.41               | 0.78       | 34.0| 57.6   | 28.0 | 84.9  |
| 5-C       | 600         | 0.41               | 0.78       | 33.3| 24.7   | 55.6 | 84.0  |
| 5-D       | 1200        | 0.41               | 0.78       | 34.4| 20.4   | 58.1 | 82.0  |
| 6-A       | 400         | 0.29               | 1.00       | 33.4| 4.8    | 57.5 | 69.8  |

### Table 1

Overview of simulation results (the type-1-error was set to 5% (two-sided), the overall sample size is 2600 patients).
controlled and preferably double-blind clinical trials. Obviously, single-arm clinical trials are also in this situation less robust and more fragile: If event rates change during the COVID-19-period, there is no control available to identify that this aspect is unspecific and affects the trial without involvement of the treatment. Lack of internal control may obviate the ability to explain unspecific effects on trial patients and cannot be distinguished from systematic effects of treatment. Suitable external controls cannot be found and historical comparisons are no longer meaningful. We assume that in these circumstances the discussion will be even more complex.

Many decisions during the recent months had to be made under pressure to safeguard the trial patients in regions where it had been possible to prepare, or to do the best for them, in regions where this was not possible. Pressure on all stakeholders has been and still is high. It is therefore of utmost importance to step back now and refrain from immediate decision making about the further conduct of a study. We recommend to carefully consider the next steps to be taken. Obviously after termination of trials all studies that had to proceed through the pandemic will require careful analysis and discussion of their interpretation, but now is the time to carefully prepare this to the extent possible.

### Funding

This work has been funded by the Federal Ministry of Education and Research (BMBF; 01KG1303).

### References

[1] Guidance on the Management of Clinical Trials during the COVID-19 (Coronavirus) pandemic, Version 3.0, (28/04/2020) [https://ec.europa.eu/health/sites/health/files/files/eudralex/vol-10/guidanceclinicaltrials_covid19_en.pdf](https://ec.europa.eu/health/sites/health/files/files/eudralex/vol-10/guidanceclinicaltrials_covid19_en.pdf).

[2] FDA Guidance on Conduct of Clinical Trials of Medical Products during COVID-19 Public Health Emergency, updated Version of 02/07/2020, (2020) [https://www.fda.gov/media/136238/download](https://www.fda.gov/media/136238/download).

[3] U. Bavendiek, D. Berliner, L.A. Dávila, et al., Rationale and design of the DIGIT-HF trial (DIGitoxin to improve ouTcomes in patients with advanced chronic heart Failure): a randomized, double-blind, placebo-controlled study, Eur. J. Heart Fail. 21 (5) (2019) 676–684, [https://doi.org/10.1002/ejhf.1452](https://doi.org/10.1002/ejhf.1452).

[4] Deutsches Ärzteblatt, COVID-19-Folgen: Weniger Herzinfarktpatienten und geringere Impfquote, (2020) [https://www.aerzteblatt.de/nachrichten/113034/COVID-19-Folgen-Weniger-Herzinfarktpatienten-und-geringere-Impfquote](https://www.aerzteblatt.de/nachrichten/113034/COVID-19-Folgen-Weniger-Herzinfarktpatienten-und-geringere-Impfquote).

[5] D. Castagno, M.C. Petrie, B. Cloggett, J. McMurray, Should we SHIFT our thinking about digoxin? Observations on ivabradine and heart rate reduction in heart failure, Eur. Heart J. 33 (2012) 1137–1141.

[6] Statistisches Bundesamt, Bundesamt, Sterbefallzahlen im April 2020 8% über dem Durchschnitt der Vorjahre, (2020) [https://www.destatis.de/DE/PressePressemeldungen/2020/05/0520194_12621.html](https://www.destatis.de/DE/PressePressemeldungen/2020/05/0520194_12621.html).

---

**Box 1**

**Main Recommendations.**

Careful stepwise decision making is essential:

1. Development of center / region / individual risk covariates which depict individual patient risk for COVID-19 is required
2. Extensive assessment of consistency of blinded overall event rates should be done
3. Discussion of unblinded event rates and treatment effects should be only considered in case of evidence for heterogeneous effects and should only be done by the responsible DSMB
4. “Do nothing” is in most studies the best option to ensure interpretability of the study results we respect to the impact of COVID-19 and the overall benefit/risk assessment