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.
Impact of COVID-19 pandemic on oncology clinical trial design, data collection and analysis

Rui (Sammi) Tang\textsuperscript{a}, Jian Zhu\textsuperscript{a}, Tai-Tsang Chen\textsuperscript{b,}\*, Fang Liu\textsuperscript{c}, Xun Jiang\textsuperscript{d}, Bo Huang\textsuperscript{e}, J. Jack Lee\textsuperscript{f}, Robert A. Beckman\textsuperscript{g}

\textsuperscript{a} Servier Pharmaceuticals, Boston, MA, USA
\textsuperscript{b} GlaxoSmithKline, Collegeville, PA, USA
\textsuperscript{c} Merck & Co., Inc., Kenilworth, NJ, USA
\textsuperscript{d} Amgen Inc., Thousand Oaks, CA, USA
\textsuperscript{e} Pfizer Inc., Groton, CT, USA
\textsuperscript{f} The University of Texas MD, Anderson Cancer Center, TX, USA
\textsuperscript{g} Lombardi Comprehensive Cancer Center and Innovation Center for Biomedical Informatics, Georgetown University Medical Center, Washington, DC, USA

\section{1. Introduction}

Since the coronavirus disease 2019 (COVID-19), a disease caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), was declared a pandemic by the World Health Organization in March 2020, the appalling huge numbers of infections and deaths have led to a paradigm shift and reshaped the entire health care system. While the primary focus has been on treatment and prevention of COVID-19 infection, this unprecedented global crisis also has had a profound impact on clinical trials. Multiple regulatory agencies and health authorities, including the U.S. Food and Drug Administration (FDA) and the European Medicines Agency (EMA), have issued guidance on conducting clinical trials during the pandemic \cite{1-7}, emphasizing assuring trial participant safety, maintaining compliance with GCP and minimizing the risk to trial integrity and data quality. FDA also issued specific guidance on statistical considerations \cite{8}.

Cancer patients are among the most impacted and vulnerable patient populations during the COVID-19 pandemic due to the requirement for regular on-site visits for tumor assessments and drug administrations. This is particularly the case when most clinical trials are designed to study the efficacy and safety of investigational agents in a controlled setting by synchronizing multiple activities at scheduled visits: drug...
distribution/treatment, disease assessments, and lab sample collection, to name a few. Thus, those clinical trial operation routines are likely to be disrupted by the pandemic and public health control to curb its spread, affecting patient enrollment, disease course, treatment, follow-up, and outcome evaluation. So far delays are observed in the initiation of new trials as well as recruitment pauses of new patients to ongoing studies for the safety consideration of patients.

However, the COVID-19 pandemic also interferes with the conduct of ongoing trials, especially for patients already enrolled. Fig. 1 illustrates selected scenarios of ongoing trials during the pandemic (trial initiated before the pandemic) with three possible time periods during a trial (pre-pandemic, pandemic and post-pandemic) according to different types of endpoints: pre-pandemic period refers to the period before COVID-19 impacted the trial, and post-pandemic period refers to the period when COVID-19 has no or little impact on the trial. The disruption of treatment and health care poses a serious threat to cancer patients and trial quality. The re-assessment of the benefit-risk of conducting clinical trials in cancer patients needs to consider unmet medical needs, severity and stage of the disease and whether alternative therapies are available. The public control measures may lead to quarantines, travel restrictions, site closures or limitations on physical access to sites. Radiation, chemotherapy, surgery, and stem-cell transplantation may be postponed for some cancer patients [9]. A survey [10] conducted by CRI and IQVIA indicated that some institutions reported delayed patient care, along with challenges in treatment administration and concerns over the safety of patients and staff; and only 20% and 14% of the investigators from the US and Europe respectively indicated that the enrollment was not impacted. In addition to the operational challenges, efficacy assessment of investigational cancer treatments may be impacted in several ways (Table 1).

The purpose of this article is to assess via simulation the impact of the COVID-19 pandemic specifically on oncology trials, to illustrate a quantitative risk assessment process of these COVID-19 effects, and to discuss potential mitigation strategies for study design, data collection, endpoints and analyses.

2. Methods: quantitative risk assessment of COVID-19

At the current state of the pandemic, we have yet to collect sufficient clinical trial data to help us understand the true impact to oncology clinical trials. Therefore, the use of modeling and simulation plays an important role in helping us understand and mitigate the potential impact of COVID-19 on clinical trial conduct. It is easy and intuitive to quantify the impact attributable to the factors of interest by controlling the others. With appropriate assumptions, the simulation results can be utilized to inform decisions at the design stage in anticipation of COVID-19 impact, as well as for ongoing trials to evaluate the impact and form potential mitigation strategies.

For demonstration purposes, our simulation study focused on the comparisons of OS and PFS between a treatment arm and a control (standard-of-care: SOC) arm with a 1:1 randomization ratio. In general, similar approaches can be applied to other endpoints in different designs (e.g., time to treatment failure, event-free survival). We considered two of the most commonly statistical inference methods used in oncology trials for time to event data: 2-sided log-rank test and hazard ratio (HR) estimated from the Cox regression model assuming proportional hazards. Performance measures for the statistical inference include the statistical power of the log-rank test under the alternative hypothesis that the therapy is effective, as well as bias in estimating the HR (i.e. difference between the average estimate and the true value). The impact of COVID-19 on statistical inference was evaluated through extensive simulations by varying the proportions of patients and magnitudes of COVID-19 infection/treatment interruption, region impact or missing disease assessments between the two arms. As multiple factors may impact the operating characteristics differently, for demonstration purposes, the impact of a given factor (or factors) was evaluated with only that factor (or factors) being considered in each simulation scenario. However, in practice, multiple relevant factors can be considered simultaneously in one simulation study to evaluate the joint impact.

Each simulation contained 1000 replicates and each trial was originally designed under the alternative hypothesis hazard ratio of 0.6 (40% risk reduction on therapy) for both PFS and OS (median PFS 20 months vs. 12 months; multiple scenarios of median OS were considered), 90% power, two-sided 5% false positive rate. The expected number of events in each trial was 162 (the total sample size was 200 for OS trials and 300 for PFS trials). We assessed the impact on the OS analysis due to deteriorated efficacy of each arm with or without regional differences, respectively. We also assessed the impact on PFS analysis due to missing disease assessments. Both noninformative and informative censoring schemes were investigated for missing disease assessments. The OS and PFS simulation flow charts (Fig. S1a and Fig. S1b) and additional details of the simulation set-up, including the length of the enrollment period before the pandemic, the duration of the pandemic, uniform enrollment and change point distribution as well as the simulation codes, can be found in the supplementary materials.

Motivated by the examples illustrated in Table 1, we conducted simulations to assess the impact of the pandemic (such as treatment interruption) on OS and PFS under two scenarios, i.e., balanced vs. imbalanced. Since the findings of the PFS simulations assessing the COVID impact were similar to those of OS, only the OS simulations and results were presented herein. We assumed 30%, 20% and 100% - (30 + 20) = 50% of the treatment arm patients were enrolled before, during and after the pandemic. The ratio of the proportions of patients impacted between SOC and treatment arm ranged from 0 (i.e., no impact) to 1.5. For example, if the ratio is 0.5, then 0.5*30% = 15%, 0.5*20% = 10%, and 100%-0.5*(30 + 20)% = 75% of the SOC patients were enrolled before, during and after the pandemic, for a total of 100% of the SOC patients.

To further incorporate the regional differences in multi-regional clinical trials, the simulated HR was fixed at 0.6 in low pandemic impact regions and high pandemic impact regions (the same as the HR without pandemic impact if patients enrolled after pandemic), though hazards for each arm could be different across regions. In patients enrolled before and during pandemic, we assumed 50% of patients who enrolled before or during pandemic in the control arm were from regions with low pandemic impact (another 50% patients are from regions with high pandemic impact), while the percentage of patients who enrolled before or during pandemic in treatment arm from low pandemic impact regions varied from 10%–90%. Details are included in Table S1-S4 in the supplementary materials.

In order to assess the impact of missing tumor assessments on PFS, we varied the ratio of the probability of missing each visit between two arms (treatment vs control) such that it varied from 0 to 2 (with an SOC arm missing assessment probability of 5%, so the treatment arm missing assessment probability varies from 0 to 10%). Note that the impact on the control arm is set to be fixed in PFS simulations, compared with the fixed impact on the treatment arm in OS simulations. Such imbalance in the missing pattern can be introduced for certain reasons (e.g., more patients from one arm were enrolled in regions more heavily impacted by COVID-19). We considered the case in which missing assessments are independent (e.g. due to lock-down or inconvenience). In the presence of such missing disease assessments, we evaluated the impact of imbalanced missing tumor assessments across arms on the statistical power and estimation bias of HR using two censoring rules: (1) Intent-to-Treat (ITT) considered all available PFS events regardless of the missing pattern, and (2) Consecutive-missing Censoring (CMC) that censors the PFS event after two or more missed visits [11].

For each subject in the simulated trials, we generated the underlying PFS from exponential distributions. After applying a disease assessment schedule (e.g., once every 2 months), the underlying PFS was observed or censored at one of the disease assessment visits. To focus on the impact of missing assessments due to COVID-19, no other censoring such
Fig. 1. Illustration of selected trial scenarios for different endpoints including overall survival (A, B), overall response (C, D) and fixed follow-up endpoints such as minimal residual disease status (E, F). **Fig. 1-A, 1-C, 1-E** illustrate the scenarios where enrollment was completed before the pandemic. **Fig. 1-B, 1-D, 1-F** illustrate the scenarios where enrollment was ongoing during and after the pandemic.

**TTE:** Time to Event; **ORR:** Objective/Overall Response Rate; **FPI:** First Patient In; **LPI:** Last Patient In; **DCO:** Data Cut-Off; **F/U:** Follow-up; **EP:** Evaluation Period
conclusions are similar, in this section we presented the results for the larger event size, and additional results are available in the Supplementary Materials.

3.1. Impact on OS

Fig. 3 summarizes the power and estimation bias of HR when there was imbalanced pandemic impact on OS across the two arms assuming no regional difference (left panel) and due to regional differences of pandemic impact (right panel).

For the imbalanced impact on OS assuming no regional difference, if more patients were impacted in the treatment arm (ratio < 1), the power will be reduced (Fig. 3-A), and the hazard ratio may be overestimated, i.e., estimated HR on average is greater than the true value of 0.6 (Fig. 3-C). As a result, the treatment effect is underestimated. For example, if there are substantially more patients impacted in the treatment arm, our simulation indicates a larger loss of power that can be severe enough to cause likely failure of a study that would have been sufficiently powered. If the number of patients impacted by COVID 19 in treatment arm is doubled in control arm in our simulation setting, the study power can be reduced to 50%. The power loss and HR bias are minimal when the proportion of impacted patients are the same across two arms (ratio = 1). Besides treatment impact balance, another important factor we evaluated is regional effect impact on clinical trials, For the imbalanced impact on OS due to regional differences, when the percentage of patients from low pandemic impact region are balanced across two arms (both are 50% in our cases), the power loss and HR bias are minimal.

For the imbalanced impact on OS due to regional differences, when the percentage of patients from low pandemic impact region are balanced across two arms (both are 50% in our cases), the power loss and HR bias are minimal.

as censoring due to subsequent anti-cancer therapy was assumed.

3. Results

For each simulation study, we considered event sizes corresponding to 90% and 80% under the normal circumstance. Since the results and
Fig. 3. Simulation evaluation of imbalanced impact across two arms assuming no regional difference (3-A, 3-C) and due to regional differences of pandemic impact (3-B, 3-D). Fig. 3-A and Fig. 3-C: 30%, 20% and 50% of the treatment arm patients were enrolled before, during and after the pandemic. The ratio of the proportions of patients impacted between control arm (i.e. SOC arm) and treatment arm ranged from 0 (i.e., no impact in the control arm) to 1.5; Fig. 3-B and Fig. 3-D: 50% of patients enrolled before or during pandemic in the control arm were from regions with low pandemic impact, and the percentage varied from 10%–90% in the treatment arm.
impact in the treatment arm is greater than that in the control arm (set at 50%), power will be lower than 90% (e.g., around 85% [Fig. 3-B] when 20% of the patients in the treatment arm are from low pandemic impact regions 80% from high pandemic impact regions) and HR will be overestimated (Fig. 3-D). Similar patterns were observed under the null hypothesis (Fig. S3).

In summary, more patients impacted by pandemic in the treatment arm may result in a loss of power and overestimate of the HR. Imbalanced enrollment from regions with high/low pandemic impact across two arms may also result in loss of power and bias of HR.

3.2. Impact on PFS

Fig. 4 provides the power and bias of HR estimates when the median PFS values for both arms are long (20 months for treatment arm vs 12 months for control arm) and tumor assessment is scheduled once every 2 months. If the chance of a tumor assessment being missing due to COVID-19 is low (probability for each disease assessment visit being missing set to be 5% in the SOC arm, and 0–10% in the treatment arm), the ITT rule will lead to minimal bias of HR estimation and subsequently negligible impact on power. However, analysis based on the CMC rule will be biased, e.g., the study will suffer from power loss if the chance of missing disease assessment in the treatment arm is lower than the control arm. The CMC rule is also operationally undesirable. Even with a low chance of missing the disease assessment due to COVID-19, it will require a longer follow-up to accrue the target event number due to censoring patients who would have been counted as events following the ITT rule.

We perform a case study from the above simulation to further demonstrate the impact. In this case study, we assume the probability of missing each disease assessment is 5% in the control arm, and 2% in the treatment arm (ratio = 0.4 on the X axis in Fig. 4), i.e. patients in the control arm are more likely to miss disease assessments. The average HR from the 1000 trials is 0.651 by the CMC rule with a power loss of 12% (reduced from 90% to 78%). For comparison, the average HR by the ITT rule is 0.602 with almost no power loss.

When the chance of missing each assessment increases in both arms (set to be 8% in the control arm), simulation suggests that the target event number may never be reached due to too many patients being censored by the CMC rule, which means additional patients may need to be enrolled. This may further prolong the study duration and increase cost. The magnitude of bias will also be larger. However, even when the chance of missing each assessment is 10% in the control arm (See Fig. S2 in Supplementary materials), the ITT rule still provides robust results with power between 87%–92% regardless of a higher or lower chance of missing assessments in the treatment arm. Similar patterns were observed under the null hypothesis (Fig. S4).

We also conducted additional simulations with shorter median PFS values (10 months for treatment arm vs 6 months for control arm) and various probabilities of missing assessments. The results are similar. In summary,

- Analyzing PFS using the CMC rule may introduce bias even with low probability of missing disease assessments, especially when the missingness is imbalanced across arms. The direction of the bias also depends on which arm has the higher chance of missing assessments, and therefore may lead to power loss. Apparent power gains may include false positives. The CMC rule may also extend the study duration and increase the cost due to censoring the events observed after the consecutive missing visits.
- Compared to the CMC rule, the ITT rule is much more robust against imbalanced frequency of missing disease assessments across arms, even if the overall probability of missing each assessment is large across the study.
- Even though the CMC rule may be one of the preferred alternative censoring rules for handling consecutive missing visits, based on the above evaluation, we do not recommend applying this rule on consecutive missing visits due to COVID-19.

A decision tree for choosing the appropriate methods for PFS evaluation in the presence of missing disease assessments due to COVID-19 is provided below (Fig. 5).
missing Censoring.

COVID infection. On the other hand, enhanced immunity introduced by the treatment effect). A similar outcome was observed when the proportion of patients impacted in high pandemic regions was higher in the treatment arm. For the assessment of PFS in the presence of missing assessments, the ITT analysis strategy performed better than the CMC analysis strategy.

For ongoing trials, the simulation can be modified to incorporate data already observed and predict what could happen to future data points given the impact of COVID-19. Such simulations would be highly informative to independent Data Monitoring Committee (IDMC) members in understanding the probability of success and potential bias and making appropriate recommendations to the sponsor (e.g., pause/stop trial due to COVID-19 impact), upon which the sponsor could start planning potential remedies (e.g., sensitivity analysis, temporary closure of sites, sample size adjustment, initiation of discussions with regulatory agencies). The IDMC could be a valuable resource to assess the medical and statistical risks due to COVID-19 and provide advice to the trial team. If an IDMC is already in place for a trial, the sponsor should consider revising the IDMC charter to include additional analyses related to risk assessment. However, operational firewalls such as the designation of independent statistical teams in charge of simulation using unblinded data should be properly established to maintain trial integrity and minimize operational bias.

The actual impact for a given trial would depend on features of the trial design, and thus careful and timely monitoring with modeling and simulation is critical. Also, since simulation requires certain assumptions that are hard to obtain empirically (e.g., impact on HR of COVID-19 infection), we recommend covering a broad range of plausible values in the simulation and to test the robustness or potential tipping-points of the design/data package. Furthermore, in the randomized clinical trial setting, due to the different mechanisms of the drugs and routes of administration (oral vs parenteral), the impact of COVID-19 on the comparative arms could potentially be different, and the common Cox proportional hazards model of time to event analysis may be challenged due to potential violation of the proportional hazards (PH) assumption, i.e. that the risk is constant over time. The log-rank test may suffer significant loss of power and the hazard ratio may be no longer interpretable under these circumstances. The restricted mean survival time (RMST), i.e. area under the survival curve, is a robust and clinically interpretable summary measure of the survival time distribution that does not rely on the PH assumption. Unlike the median survival time, it is estimable even under heavy censoring. There has been considerable methodological research [12, 13] on the use of RMST to estimate treatment effects as an alternative to the hazard ratio (HR) approach, and it can still have robust power under non-proportional hazards and provide clinically meaningful estimation of clinical benefit [14-16]. Furthermore, one can also consider appropriate adaptive designs to overcome the various challenges posed by the pandemic, and well utilize the role of data monitoring committees. A whole host of issues and challenges encountered in the clinical trial conduct during the pandemic and potential mediations such as the use of adaptive designs combining information across stages (e.g., pre-/ during / post-pandemic), unplanned trial modifications to respond to the pandemic, operational challenges, and the impact on the estimand and the study integrity and interpretability have also been investigated [17]. As the data gradually become available, we hope our contribution will enable us and other researchers to continue investigating the impact of COVID-19 pandemic on the oncology clinical trial conduct.

4. Conclusion

While most researchers’ primary focus has been on treatment and prevention of COVID-19 infection, this unprecedented global crisis also has had a profound impact on clinical trials. We conducted extensive simulations to identify and assess the impact of the COVID-19 pandemic specifically on oncology trials and discuss potential mitigation strategies for study design, data collection, various endpoints, and analyses. In general, one would hope that a randomized double-blind study design would minimize the impact of the imbalance between arms. Given a vast number of cancer treatments with various mechanisms of action, individuals with immunosuppression, i.e., a side effect of most cancer therapies, are more susceptible to severe morbidity and mortality from COVID infection. On the other hand, enhanced immunity introduced by immune checkpoint inhibitors may be expected to improve protection against COVID. In our simulation study, we have also investigated an imbalanced COVID-19 impact on the study.

An imbalance that was in favour of the control arm would lead to a loss of power and overestimation of the hazard ratio (underestimation of the treatment effect). A similar outcome was observed when the proportion of patients impacted in high pandemic regions was higher in the treatment arm. For the assessment of PFS in the presence of missing assessments, the ITT analysis strategy performed better than the CMC analysis strategy.

Fig. 5. Decision tree for choosing the analysis method and trial adjustment to handle missing disease assessments. ITT: Intent-To-Treat; CMC: Consecutive-

14, 15]}. Further-
Role of the funder

Not applicable.

Disclosures

The authors have no conflicts of interest to report.

Prior presentation

The content of this manuscript has not been presented elsewhere.

Author contributions

All authors participated in the discussions, contributed to the writing of the initial drafts; reviewed and approved the final manuscript.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cct.2022.106736.

References

[1] FDA Guidance on Conduct of Clinical Trials of Medical Products during COVID-19 Pandemic: Guidance for Industry, Investigators, and Institutional Review Boards. https://www.fda.gov/media/136238/download, 2020.
[2] EMA Guidance on the Management of Clinical Trials during the COVID-19 (Coronavirus) pandemic; Version 2. https://ec.europa.eu/health/sites/health/files/eudralex/vol-10/guidanceclincialtrials_covid19_en.pdf, March 27, 2020.
[3] EUROPE: COVID-19 Overview of European Regulatory Authority Guidance; Edition 1.0, March 17, 2020.
[4] EMA Points to consider on implications of Coronavirus disease 5 (COVID-19) on methodological aspects of ongoing clinical 6 trials: Draft. https://www.ema.europa.eu/en/documents/scientific-guideline/points-consider-implications-coronavirus-disease-2019-methodological-aspects-ongoing-clinical-en.pdf, March 25, 2020.
[5] PMDA Q&A on how to manage clinical trials during the COVID-19 pandemic. http://www.pmda.go.jp/files/000235164.pdf, March 27, 2020.
[6] NHS HRA Guidance about COVID-19 for sponsors, sites and researchers; v2.2. https://www.hra.nhs.uk/covid-19-research/covid-19-guidance-sponsors-sites-and-researchers/, March 26, 2020.
[7] MIHA, Guidance on Managing Clinical Trials during Coronavirus (COVID-19). https://www.gov.uk/guidance/managing-clinical-trials-during-coronavirus-covid-19, March 19 2020.
[8] FDA Guidance on Statistical Considerations for Clinical Trials During the COVID-19 Public Health Emergency. https://www.fda.gov/media/139145/download, June 2020.
[9] ASCO, COVID-19 Patient Care Information. https://www.asco.org/asco-coronavirus-information/care-individuals-cancer-during-covid-19.
[10] Impact of COVID-19 on oncology clinical trials – Nature reviews drug discovery. https://www.nature.com/articles/d41573-020-00093-1.
[11] J.S. Denne, A.M. Stone, R. Bailey-Iacona, T.T. Chen, Missing data and censoring in the analysis of progression-free survival in oncology clinical trials, J. Biopharm. Stat. 23 (5) (2013) 951–970, https://doi.org/10.1080/10543406.2013.813515 (PMID: 23957509).
[12] P. Royston, M. Parmar, Restricted mean survival time: an alternative to the hazard ratio for the design and analysis of randomized trials with a time-to-event outcome, BMC Med. Res. Methodol. 13 (1) (2013) 152–166.
[13] H. Uno, B. Chaggett, L. Tian, et al., Moving beyond the hazard ratio in quantifying the between-group difference in survival analysis, J. Clin. Oncol. 32 (22) (2014) 2380–2385.
[14] L. Tian, H. Fu, S.J. Ruberg, H. Uno, L.J. Wei, Efficiency of two sample tests via the restricted mean survival time for analyzing event time observations, Biometrics. 74 (2) (2018 Jun) 694–702.
[15] B. Huang, P.F. Kuan, Comparison of the restricted mean survival time with the hazard ratio in superiority trials with a time-to-event endpoint, Pharm. Stat. 17 (3) (2018 May) 202–213.
[16] B. Huang, L.J. Wei, E.B. Ludmir, Estimating treatment effect as the primary analysis in a comparative study: moving beyond P value, J. Clin. Oncol. JCO-19 (2020 Apr).
[17] T. Hamasaki, F. Brez, F. Cooner, L.M. LaVange, M. Posch (Eds.), Stat. Biopharmaceut. Res. 12 (4) (2020) [Special Issue].