Consequences of Performing Parallel Dose Finding Trials in Heterogeneous Groups of Patients

Bethany Jablonski Horton, John O'Quigley, Mark R Conaway

JNCI Cancer Spectrum (2018) 3(2): pkz013

doi: 10.1093/jncics/pkz013
First published online May 7, 2019
Commentary

COMMENTARY

Consequences of Performing Parallel Dose Finding Trials in Heterogeneous Groups of Patients

Bethany Jablonski Horton, John O'Quigley, Mark R. Conaway

See the Notes section for the full list of authors' affiliations.
Correspondence to: Bethany Jablonski Horton, PhD, Department of Public Health Sciences, University of Virginia, P.O. Box 800717, Charlottesville, VA 22908 (e-mail: bhorton@virginia.edu).

Abstract

Patient heterogeneity, in which patients can be grouped by risk of toxicity, is a design challenge in early phase dose finding trials. Carrying out independent trials for each group is a readily available approach for dose finding. However, this often leads to dose recommendations that violate the known order of toxicity risk by group, or reversals in dose recommendation. In this manuscript, trials for partially ordered groups are simulated using four approaches: independent parallel trials using the continual reassessment method (CRM), Bayesian optimal interval design, and 3 + 3 methods, as well as CRM for partially ordered groups. Multiple group order structures are considered, allowing for varying amounts of group frailty order information. These simulations find that parallel trials in the presence of partially ordered groups display a high frequency of trials resulting in reversals. Reversals occur when dose recommendations do not follow known order of toxicity risk by group, such as recommending a higher dose level in a group of patients known to have a higher risk of toxicity. CRM for partially ordered groups eliminates the issue of reversals, and simulation results indicate improved frequency of maximum tolerated dose selection as well as treating a greater proportion of trial patients at this dose compared with parallel trials. When information is available on differences in toxicity risk by patient subgroup, methods designed to account for known group ordering should be considered to avoid reversals in dose recommendations and improve operating characteristics.

This work is motivated by existing dose finding clinical trials conducted in groups of patients, where some information is available that the risk of toxicity may differ by group. Carrying out separate, independent trials for each group (parallel trials) will often be a suboptimal way to proceed. Sharing information between groups will often lead to improved accuracy. Many examples of parallel trials are found in the literature. LoRusso et al. (1) and Ramanathan et al. (2) both implemented trials for four groups, where groups were defined by normal, mild, moderate, and severe liver dysfunction, and used the 3 + 3 design separately within each group. These are examples of trials with the greatest amount of group order information, where the groups are completely ordered. For example, with a given dose level, it is expected that a greater risk of toxicity among patients is associated with a higher level of liver dysfunction. This suggests that a group with greater liver dysfunction should not have a maximum tolerated dose (MTD), which is a greater dose level than that of a group with less liver dysfunction. Most commonly, in practice, parallel trials are implemented in which patients in one group are allocated to doses without regard to the toxicity information available from patients in other groups. Hence, group order information is ignored, and it is possible to observe MTD selection that does not preserve the known group order.

Use of independent parallel trials is a readily available method for identifying the MTD for each group, but independent trials ignore the available group order information. Failure to account for group order structure creates issues that should be considered more closely. The primary concern with the use of parallel trials by group is in identifying group-specific MTDs that do not follow the known group frailty order. Suppose a particular group is known to be more frail; thus, the MTD should be no greater than that of other groups considered. It is problematic if the final MTD selection for this group is a higher dose

Received: November 8, 2018; **Revised:** February 1, 2019; **Accepted:** March 18, 2019

level than other groups considered, which are known to be less frail. Three methods have been published recently that account for either partial or complete group order information in dose finding trials (3–5). Because the group order information is not ignored in the methods designed for groups, it is not possible to make MTD recommendations that do not preserve the known group order structure. In addition, operating characteristics indicate improved dose selection when using dose finding methods designed to incorporate group order information.

The aim of this manuscript is to provide information regarding concerns with implementation of independent parallel trials for dose finding in the presence of groups and to offer comparison of operating characteristics when using a method that appropriately accounts for the known group order. In this manuscript, we will first describe the problems associated with implementation of parallel trials for groups, followed by a description of the simulations considered and discussion of the simulation results.

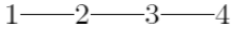
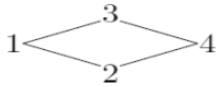
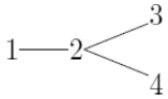
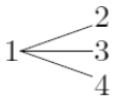
Group Order Structures and Reversals

By our definition, a reversal occurs when MTD selection violates the known group order. The type and number of possible reversals depends greatly on the amount of information that is known about the group frailty order at the beginning of the trial. In the presence of four groups, there are four group ordering structures considered in the presented simulations: complete order, loop order, twig order, and the simple tree. In this commentary, it is assumed that higher group numbers are associated with greater frailty, that is, an increased susceptibility at each level to suffering a dose limiting toxicity (DLT). The MTD for a group known to be more frail should be no greater than the MTD of a group that is less frail. Table 1 presents the known information for the group orders considered.

Complete Order

The complete order, shown in Table 1, offers the most information about the relationships between groups. Therefore, $MTD_{G1} \geq MTD_{G2} \geq MTD_{G3} \geq MTD_{G4}$ is known with no unknown relationships between groups. Due to the level of information, this structure also allows for the greatest number of reversals to be observed within a trial. Reversals can be observed between all pairs of groups, with the potential of observing six reversals within a trial.

Table 1. Types of group order

Group order	Level of known information	Least frail ⟷ Most frail
Complete order	Most	
Loop order	Intermediate	
Twig order	Intermediate	
Simple tree	Least	

Loop Order

The loop order, shown in Table 1, has an intermediate level of information about the relationship between groups. Here, it is known that $MTD_{G1} \geq MTD_{G2} \geq MTD_{G4}$ that and $MTD_{G1} \geq MTD_{G3} \geq MTD_{G4}$. In this case, it is not possible to specify the relationship between MTD_{G2} and MTD_{G3}, creating a partial order among these groups. Because there is less information known compared with the complete order case, it is possible to observe five reversals for partial orders of this type. Reversals can occur between groups 1 and 2, 1 and 3, 1 and 4, 2 and 4, and between groups 3 and 4. No reversal can exist between groups 2 and 3 because this order relationship is unknown.

Twig Order

The "twig order," shown in Table 1, is also considered to have an intermediate level of information with regard to the relationships between groups. The known information for the relationship between groups includes $MTD_{G1} \geq MTD_{G2} \geq MTD_{G3}$ and $MTD_{G1} \geq MTD_{G2} \geq MTD_{G4}$, whereas the relationship between MTD_{G3} and MTD_{G4} cannot be specified. With this partial order structure, it is possible to observe the following five reversals: between groups 1 and 2, 1 and 3, 1 and 4, 2 and 3, and between groups 2 and 4. There can be no reversal between groups 3 and 4 because this order relationship is unknown.

Simple Tree

The simple tree, shown in Table 1, is a partial order that contains the least amount of information among order structures considered. Three group order relationships are known: $MTD_{G1} \geq MTD_{G2}$, $MTD_{G1} \geq MTD_{G3}$, and $MTD_{G1} \geq MTD_{G4}$. Only three types of reversals are possible: between groups 1 and 2, 1 and 3, and 1 and 4. It is not possible to specify the relationships between MTD_{G2}, MTD_{G3}, and MTD_{G4}; thus, reversals between these groups are not possible.

Dose Finding Methods Considered

Two approaches to dose finding trials are considered in the simulations: parallel trials and continual reassessment method for partially ordered groups (CRM-POG), the method given by Horton et al. (5) that incorporates information on the group order structure. Parallel trials are implemented using several standard dose finding methods.

Parallel Trials

Parallel trials using three dose finding methods are considered in simulations: CRM, Bayesian optimal interval design (BOIN), and the standard 3 + 3. CRM, given by O'Quigley et al. (6), is a widely recognized model-based method for dose finding. Simulations presented in this manuscript utilize the two-stage likelihood version of CRM, given by O'Quigley and Shen (7), specifying a one-parameter working model for the probability of toxicity. Simulation results for the CRM design were obtained using the "crmsim" function in R package "dfcrm," specifying the "mle" method and "empiric" model with cohorts of size 1. The skeleton was specified using the "getprior" function within the "dfcrm" package, specifying a halfwidth of 0.06 and prior guess of MTD of dose level 3.

BOIN, given by Liu and Yuan (8), allocates patients based on the specification of upper and lower cut-off values. Simulation results for the BOIN design were obtained using the "get.oc" function in R package "BOIN," specifying cohorts of size 3, and default values for all specifications. Simulation results for the 3+3 method were obtained using the "ssim3p3" function in R package "UBCRM."

Optimal Benchmark

An optimal benchmark measure is included to evaluate each method's performance in dose selection. The optimal benchmark is a nonparametric design that can be seen as a gold standard for dose finding studies (9). Operating characteristics for the optimal benchmark provide guidance and a greater ability to gauge the performance of parallel trials. When considering percentage of correct selection (PCS) and accuracy index (AI), the benchmark provides an upper bound upon which improvement is unlikely.

Alternative to Parallel Trials: CRM-POG

Simulation results presented consider CRM-POG, the method proposed by Horton et al. (5), which extends the existing CRM shift model for two ordered groups (10). Implementation of the CRM-POG is carried out in two stages. The first stage is rule based, using cohorts of size 1. Once heterogeneity is observed, at least one DLT and non-DLT, the second stage begins where dose allocation is based on estimation from a working model. To implement the second stage, a set of skeletons is developed using information on known group frailty order. The motivating example consists of trials for four groups that are completely ordered. This suggests that among the skeletons considered, the probability of toxicity at a specified dose level for group 3 is always less than or equal to the probability of toxicity for the same dose level for group 4. The skeleton maximizing the binomial likelihood is selected and used for the next dose recommendation.

This approach is computationally accessible and allows for both complete and partially ordered group information. Available R code for CRM in combination with a set of reasonable skeletons based on known order information can be used to implement the CRM-POG method. Because MTD selection follows the known group ordering, reversals cannot occur. Improvements in operating characteristics increase as the level of information increases. Group ordering structures considered are described in Table 1.

Simulation

In this section, we present the design of the simulations and discuss the simulation results. In addition to CRM-POG, independent parallel trials are simulated using the CRM, BOIN, and 3+3 methods. As a comparison, the optimal benchmark is also included (9). In all simulations, the target toxicity rate is $\theta = 0.3$, and 1000 simulated trials were generated. Nine dose-toxicity curves are considered with four dose levels, shown in Figure 1. Although patients in group 4 are considered to be the most frail, their probability of toxicity at a specific dose level should be greater than or equal to the probability of toxicity for patients in group 1, which is known for all group orders considered. Additional group order relationships depend on the known group order structure, displayed in Table 1. The dose-toxicity curves considered allow for a variety of order of severity between groups and spacing of the true MTD for each group. Scenarios in Figure 1 allow for assuming any of the four group order structures when using CRM-POG. As the amount of group order information varies, each of these four order structures is considered, and results for CRM-POG are summarized by the assumption of each type of order structure described in Table 1.

Sample sizes of 24 per group were used in simulations, chosen to equal the maximum number of patients that could be enrolled in the 3+3 trials. Although the sample size per group is fixed in these simulations, it would be reasonable to assume that in practice the number of patients in each group may vary. The greatest impact from varying group sample sizes would be seen if a 3+3 method were used, because this method is inflexible in the maximum sample that may be needed for implementation. When CRM and BOIN are used in parallel trials, having a greatly reduced sample size in one group is anticipated to have a sizable impact on operating characteristics and the ability to identify the best dose level for that group. The smallest impact would be seen when considering CRM-POG because constraints placed on MTD selection by use of the skeletons allow for dose recommendations for each group in the presence of all evaluated patients.

Summary measures in comparing parallel trials include proportion of trials experiencing at least one reversal and magnitude of reversals. Existence of a reversal indicates that MTD selection does not follow the known group frailty order, and the magnitude indicates the degree to which this is happening. For example, if at the end of parallel trials, the MTDs for a less frail and more frail group are identified as dose levels 1 and 3, respectively, a reversal has occurred with a magnitude of two dose levels. The greater the dose level discrepancy, the greater the concern of inappropriate dose selection because the group order is known and established before the implementation of the parallel trials.

Although the presence of reversals is of greatest concern with regard to running parallel trials, other operating characteristics are also considered. PCS and AI measures allow for assessing the ability of the design to address the primary aims of an early phase dose finding trial: (1) selecting the best dose available at the end of a trial, and (2) appropriately allocating patients throughout the trial. PCS gives the percentage of trials that select the dose closest to the target. CRM-POG, outlined by Horton et al. (5), is used to compare PCS and AI with parallel trials. PCS is defined as the proportion of simulations selecting the correct dose as MTD. Accuracy index, outlined by Cheung (11), was calculated for dose selection, as used in Horton et al. (12) Because reversals are not possible with CRM-POG, only PCS and AI are used to compare these methods to parallel trials.

Results

Reversals. The proportion of trials experiencing reversals depends on many factors, including the amount of known information at the beginning of the study. As shown in Table 1, the most information available is observed when complete ordering between groups is known and the least amount of information is observed with the simple tree group order structure, where the only known information is that group 1 is the least frail and MTD selection for other groups should be no greater than the MTD selection for group 1.

Figure 2 displays the proportion of simulations with at least one reversal. Because there are more opportunities for reversals in orderings with greater information, a greater proportion of trials results in at least one reversal when a greater level of group information is known. For each group order setting, the

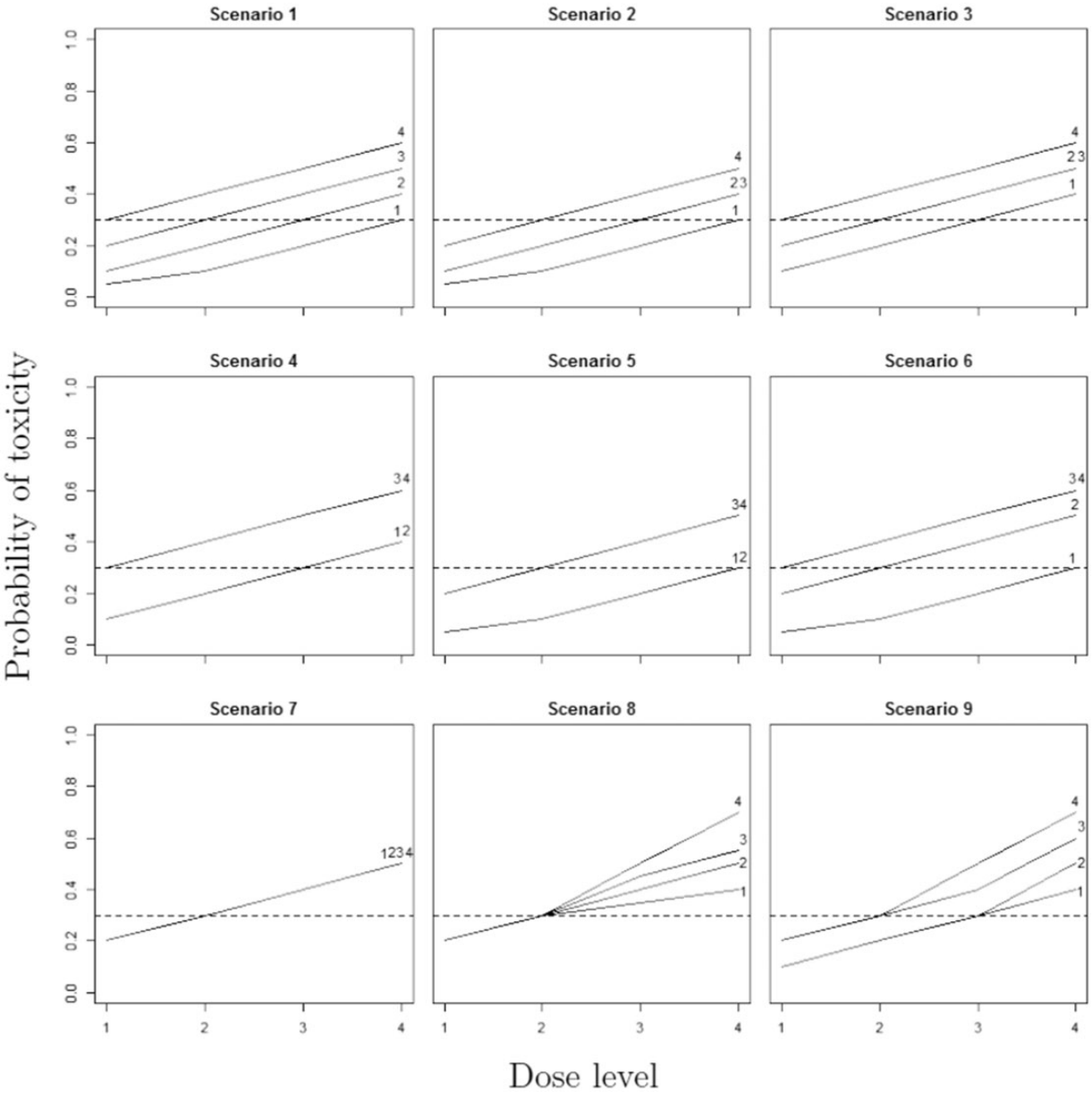


Figure 1. Dose-toxicity curves.

lowest average proportion of reversals is observed for the optimal benchmark, followed by CRM, BOIN, and 3 + 3. To illustrate the depth of this problem, consider the loop order, which contains an intermediate amount of information provided by the known group order. Here, 46% of trials using the optimal benchmark result in observing at least one reversal. This is greatly concerning considering that the benchmark is seen as a method to assess the lowest expected proportion of reversals with which to compare early phase designs. With this group order structure, 61% of 3 + 3 parallel trials result in at least one reversal.

Supplementary Figure 1 (available online) displays the average magnitude of reversals observed across scenarios considered. Note that among simulations with reversals, the distribution of reversal magnitude is weighted more toward smaller magnitude in the optimal benchmark, followed by CRM, BOIN, and last the 3 + 3.

Several conclusions can be drawn from considering the results presented in Figure 2 and Supplementary Figure 1 (available online) regarding the level of information known of the group order and the design used to implement parallel trials. When more information is known about group order (such as the complete order case), there is greater opportunity to observe a reversal. With more group order information, the magnitude of reversal also tends to be slightly greater than in cases with less group order information (such as the simple tree). Frequency of observing reversals is worst when using the 3 + 3 design, followed by BOIN, then CRM. As the optimal benchmark provides a reference value for operating characteristics to compare with dose selection designs, it is not anticipated that improvement beyond the benchmark (in this case, lower rates of reversals) will be observed using parallel trials. The percentage of simulations resulting in at least one reversal, averaged across scenarios, ranges from 27% to 54% for the optimal benchmark,

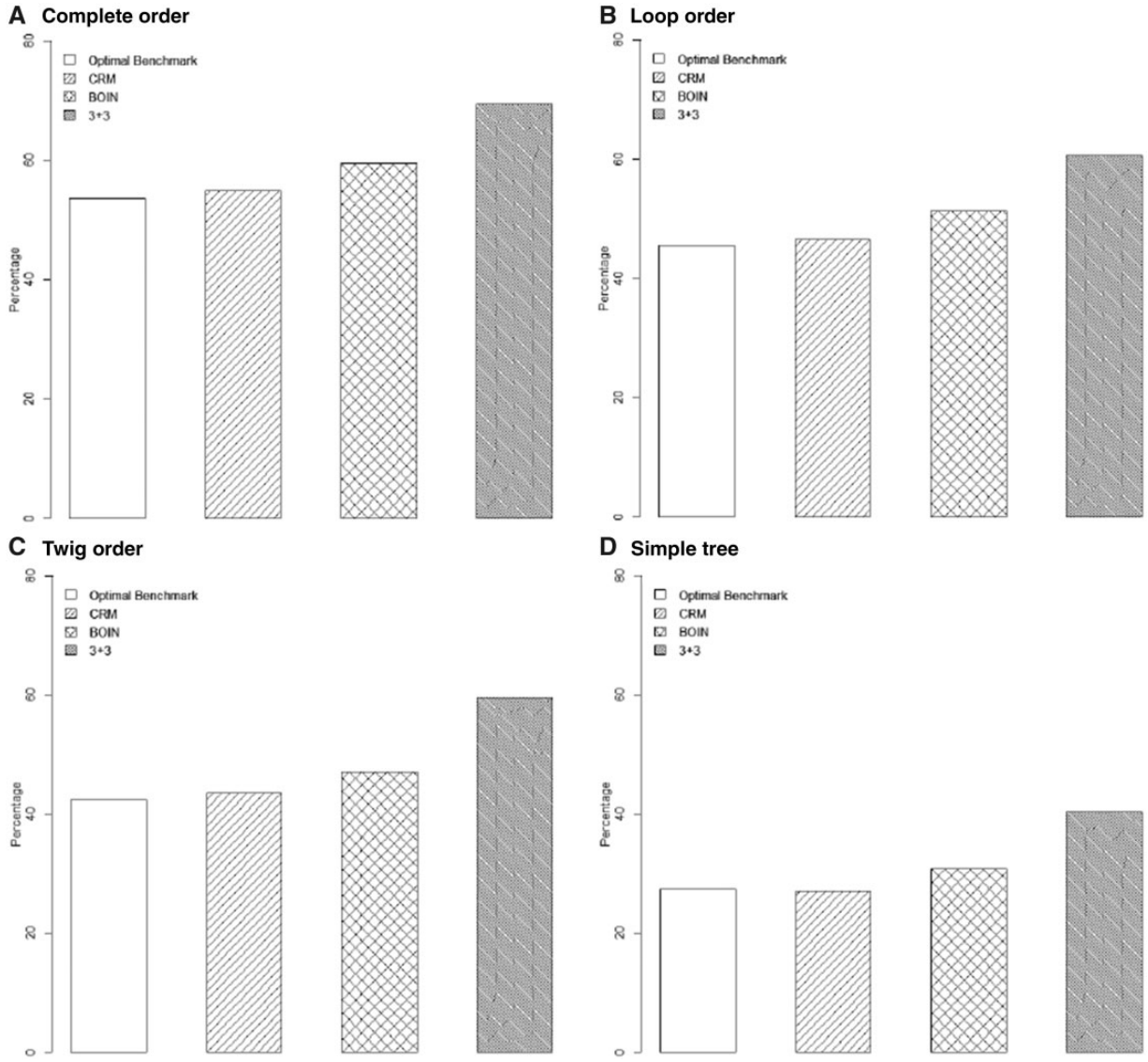


Figure 2. Parallel trials: percentage of simulations with at least one reversal. BOIN = Bayesian optimal interval design; CRM = continual reassessment method.

depending on the level of known group order information. This suggests that when there is minimal information known, in the simple tree setting, the benchmark observes a high rate of trials resulting in reversals (27%). When group orders are fully defined, in the complete order setting, 54% of parallel trials result in at least one reversal. These figures are concerning given that it is anticipated that early phase designs will either meet or exceed these rates. For instance, in the case of the $3+3$ design, the rate of observing at least one reversal ranges from 40% to 69%, depending of level of known group order information.

The concerns with frequency of observing at least one reversal and the magnitude of reversals leads to a greater question: what does MTD selection mean when using parallel trials and should there be additional plans in place to realign MTD selection? With each of the group order structures considered, there is at least some information available in comparing the level of frailty between groups. It seems unrealistic to move forward with an MTD recommendation that is a higher dose level in a group that is known to be more frail.

PCS and AI

In these simulations, the target DLT rate is $\theta = 0.3$. AI, given by Cheung (11), is used to summarize dose selection. AI considers the frequency of selecting all available doses while penalizing selection of doses with probability of toxicity further from the true MTD. Calculation of AI for dose selection is described in Horton et al. (12) using the absolute difference to measure the discrepancy between the probability of toxicity at a specified dose level and the target DLT rate.

Table 2 gives PCS and AI for all methods considered, averaged across groups and across the nine scenarios considered in the simulations. The methods considered include CRM-POG (for each order structure considered) and parallel trials using the optimal benchmark, CRM, BOIN, and the $3+3$. Results for all summary measures indicate that using the CRM-POG, with even minimal information from the simple tree structure, allows for improvement in dose selection compared with all parallel trial designs considered. The level of improvement over parallel

Table 2. Summary measures comparing trial approaches*

	Operating characteristics		Operating characteristics, reversals removed†	
Trial approach	PCS	AI	PCS	AI
CRM-POG by assumed group order structure‡				
Complete order	0.57	0.57	0.57	0.57
Loop order	0.55	0.55	0.55	0.55
Twig order	0.57	0.57	0.57	0.57
Simple tree	0.54	0.54	0.54	0.54
Parallel trials				
Optimal benchmark	0.51	0.5	0.51	0.51
CRM	0.49	0.47	0.51	0.51
BOIN	0.47	0.46	0.46	0.47
3 + 3	0.26	0.22	0.19	0.14

*Operating characteristics are averaged across groups and scenarios. AI = accuracy index for dose recommendations; BOIN = Bayesian optimal interval design; CRM = continual reassessment method; CRM-POG = continual reassessment method for partially ordered groups; PCS = percentage of correct selection.
†See Naive Approach to Parallel Trials for a description of the approach to remove reversals from maximum tolerated dose recommendations.
‡The CRM-POG method does not allow for reversals to occur; thus, there is no difference in operating characteristics once reversals are removed.

trials increases as the level of known group order information increases. For greater detail, Supplementary Table 1 (available online) gives the operating characteristics by scenario.

Although it is established that 3 + 3 greatly underperforms compared with methods such as CRM and BOIN in single group trials (13–15), this underperformance appears to be magnified in the presence of parallel trials.

Naive Approach to Parallel Trials

A reversal occurs when MTD selection does not follow the known group frailty order. For example, it would be counterintuitive to make a final MTD recommendation in a more frail group at a higher dose level than the MTD recommended for a group known to be less frail. In this case, a conservative ad hoc approach considered here is to reduce the MTD for a more frail group to not exceed the MTD of groups known to be less frail. Consider a trial with four groups and a known loop order; patients in group 1 are the least frail and patients in group 4 are the most frail. Patient frailty for groups 2 and 3 fall between groups 1 and 4, but the relationship between groups 2 and 3 cannot be specified. This suggests that the MTD for groups 2 and 3 should be no higher than that of group 1, and the MTD for group 4 should be no higher than that of all other groups. Supplementary Table 3 (available online) presents a parallel trial example. In this example, there are two reversals observed: between groups 1 and 3 and between groups 2 and 4. Both reversals have a magnitude of one dose level. Using the naive approach to "fix" the reversal so that the known group frailty order is preserved with MTD recommendations, the MTD recommendation in groups 3 and 4 are changed. This ad hoc approach was used in simulations to assess the impact of this augmentation on operating characteristics, given in Table 2. In these results, we can see that there is virtually no change in PCS and AI for both the optimal benchmark and BOIN. Marginal improvement is observed when the ad hoc approach is applied to CRM dose recommendations. There is a greater difference seen for parallel 3 + 3 trials, where the ad hoc approach leads to worse operating characteristics. Supplementary Table 2 (available online) provides additional detail by presenting the operating characteristics by scenario.

Although this ad hoc approach is considered a conservative way of handling reversals occurring in parallel trials, a better approach would be to use a method designed to incorporate known group order and prevent the possibility of a reversal while improving operating characteristics.

Discussion

This manuscript highlights an important issue of dose finding in the presence of groups when parallel trials are used. The primary concern with parallel trials is the common occurrence of reversals, where MTD selection does not follow the known group frailty order. Although parallel trials have been an accessible method of implementing dose finding trials in the presence of groups, three methods have recently been published that allow for dose finding with complete or partially ordered groups where group frailty order is always respected (3–5). Early phase dose finding trials tend to be small in size, and the methods for partially ordered groups use all information by incorporating available information on group order. These methods, which are designed for groups, not only eliminate the concern of reversals in MTD selection but also show vast improvement in operating characteristics associated with dose selection.

Several existing dose finding methods were compared in the setting of independent parallel trials to assess operating characteristics in the presence of groups, including CRM, BOIN, and the standard 3 + 3. Operating characteristics considered in comparing methods include the proportion of trials with at least one reversal, magnitude of reversals, PCS, and AI for dose selection. The optimal benchmark is used to give a lower bound on occurrence of reversals and an upper bound on PCS and AI. As has been shown previously, on average, CRM outperforms among the parallel trials, followed by BOIN and 3 + 3. The CRM-POG method proposed by Horton et al. (5) outperforms all parallel trial approaches, and the degree of improvement depends on the amount of group frailty order information that is known at the beginning of the trial.

Because reversals would need to be fixed, a conservative ad hoc approach was considered to "fix" reversals such that MTD selection could be augmented to conform to the known group frailty order. By reducing the MTD in more frail groups to not exceed the MTD of less frail groups, reversals were eliminated. Although there was minimal impact on operating characteristics for CRM and BOIN parallel trials, a sharp decrease in PCS and AI was observed for 3 + 3 parallel trials.

The primary goals of early phase dose finding trials are to select the best dose to move forward to future trials and to allocate subjects to the most appropriate dose within the current clinical trial. This work has focused on the idea of reversals in MTD selection at the conclusion of parallel clinical trials, which are a concern for the first aim of selecting the best dose to move forward to future trials. To address the second aim to allocate subjects to the most appropriate dose within the current clinical trial, within-trial reversals are an important topic that also deserves attention. Within-trial reversals occur when dose assignments are counterintuitive based on doses being assigned between groups. For example, if a patient in a group that is more frail is assigned to a dose level that is not considered to be

safe for a less frail group, this would be an instance where a counterintuitive dosing decision is made. The restrictions used in CRM-POG will not allow for within-trial reversals. Although within-trial reversals were not a focus of the current work, this is an important issue to consider in future work.

In conclusion, this manuscript demonstrates the problem of reversals in MTD selection when independent parallel trials are used in the presence of completely or partially ordered groups. In light of the high rate of reversals and other diminished operating characteristics for parallel trials, a dose finding design meant for groups should be used when any level of information regarding group frailty order (either partial or complete order) is available.

Funding

This work was supported by the National Cancer Institute at the National Institutes of Health (R01 CA142859 to MRC and BJH). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Notes

Affiliations of authors: Division of Translational Research and Applied Statistics, Department of Public Health Sciences, University of Virginia, Charlottesville, VA (BJH, MRC); Universite' Paris VI, Paris, France (JO).

We would like to thank the reviewers for their thoughtful comments in the process of reviewing this manuscript. Their insight added attention to important details in this work.

References

1. LoRusso PM, Venkatakrishnan K, Ramanathan RK, et al. Pharmacokinetics and safety of bortezomib in patients with advanced malignancies and varying degrees of liver dysfunction: Phase I NCI Organ Dysfunction Working Group Study NCI-6432. *Clin Cancer Res*. 2012;18(10):2954–2963. doi: 10.1158/1078-0432.CCR-11-2873.
2. Ramanathan RK, Egorin MJ, Takimoto CH, et al. Phase I and pharmacokinetic study of imatinib mesylate in patients with advanced malignancies and varying degrees of liver dysfunction: a study by the National Cancer institute Organ Dysfunction Working Group. *JCO*. 2008;26(4):563–569.
3. Conaway M. Isotonic designs for phase I trials in partially ordered groups. *Clinical Trials*. 2017;14(5):491–498.
4. Conaway MR. A design for phase I trials in completely or partially ordered groups. *Stat Med*. 2017;36(15):2323–2332.
5. Horton BJ, Wages NA, Conaway MR. Shift models for dose-finding in partially ordered groups. *Clin Trials*. 2019;16(1):32–40.
6. O'Quigley J, Pepe M, Fisher L. Continual reassessment method: a practical design for phase 1 clinical trials in cancer. *Biometrics*. 1990;46(1):33–48.
7. O'Quigley J, Shen LZ. Continual reassessment method: a likelihood approach. *Biometrics*. 1996;52(2):673–684.
8. Liu S, Yuan Y. Bayesian optimal interval designs for phase I clinical trials. *J R Stat Soc C*. 2015;64(3):507–523.
9. O'Quigley J, Paoletti X, Maccario J. Non-parametric optimal design in dose finding studies. *Biostatistics*. 2002;3(1):51–56.
10. O'Quigley J, Iasonos A. Bridging solutions in dose-finding problems. *Stat Biopharm Res*. 2014;6(2):185–197.
11. Cheung YK. *Dose Finding by the Continual Reassessment Method*. New York: Chapman and Hall/CRC; 2011.
12. Horton BJ, Wages NA, Conaway MR. Performance of toxicity probability interval based designs in contrast to the continual reassessment method. *Statist Med*. 2017;36(2):291–300.
13. Jaki T, Clive S, Weir CJ. Principles of dose finding studies in cancer: a comparison of trial designs. *Cancer Chemother Pharmacol*. 2013;71(5):1107–1114.
14. Paoletti X, Ezzalfani M, Le Tourneau C. Statistical controversies in clinical research: requiem for the 3 + 3 design for phase I trials. *Ann Oncol*. 2015;26(9):1808–1812.
15. Iasonos A, Wilton AS, Riedel ER, Seshan VE, Spriggs DR. A comprehensive comparison of the continual reassessment method to the standard 3+3 dose escalation scheme in phase I dose-finding studies. *Clinical Trials*. 2008;5(5):465–477.