The Journal of Infectious Diseases

MAJOR ARTICLE

Design of a Randomized Controlled Trial for Ebola Virus Disease Medical Countermeasures: PREVAIL II, the Ebola MCM Study

Lori E. Dodd,[1] Michael A. Proschan,[1] Jacqueline Neuhaus,[3] Joseph S. Koopmeiners,[3] James Neaton,[3] John D. Beigel,[2] Kevin Barrett,[1] Henry Clifford Lane,[1] and Richard T. Davey Jr[1]

[1]National Institute of Allergy and Infectious Diseases, Bethesda, and [2]Leidos Biomedical Research, Frederick National Laboratory for Cancer Research, Maryland; and [3]Division of Biostatistics, School of Public Health, University of Minnesota, Minneapolis

Background. Unique challenges posed by emerging infectious diseases often expose inadequacies in the conventional phased investigational therapeutic development paradigm. The recent Ebola outbreak in West Africa presents a critical case-study highlighting barriers to faster development. During the outbreak, clinical trials were implemented with unprecedented speed. Yet, in most cases, this fast-tracked approach proved too slow for the rapidly evolving epidemic. Controversy abounded as to the most appropriate study designs to yield safety and efficacy data, potentially causing delays in pivotal studies. Preparation for research during future outbreaks may require acceptance of a paradigm that circumvents, accelerates, or reorders traditional phases, without losing sight of the traditional benchmarks by which drug candidates must be assessed for activity, safety and efficacy.

Methods. We present the design of an adaptive, parent protocol, ongoing in West Africa until January 2016. The exigent circumstances of the outbreak and limited prior clinical experience with experimental treatments, led to more direct bridging from preclinical studies to human trials than the conventional paradigm would typically have sanctioned, and required considerable design flexibility.

Results. Preliminary evaluation of the "barely Bayesian" design was provided through computer simulation studies. The understanding and public discussion of the study design will help its future implementation.

Keywords. emerging infectious diseases; Ebola virus disease; clinical trials; adaptive design; Bayesian design.

The unprecedented Ebola virus disease (EVD) outbreak in West Africa presented enormous challenges for the global health community. Among the many challenges was the desperate need for effective therapeutics. At the onset of the epidemic, candidate therapies were in the early stages of development, but little was known about their potential clinical benefit [1]. None had been evaluated in clinical trials against EVD, and the ethics of such trials were debated [2]. Many unproven therapies were administered under a "compassionate use" policy, and those outcomes have added very little to our understanding of potential efficacy.

In the United States, for example, 10 patients with EVD received a variety of investigational therapies, either alone or in combination. Two of those 10 patients (20%) died, a mortality rate far lower than that reported in most studies from West Africa [3, 4]. An unanswered question is whether this lower mortality rate was due to better background supportive therapy (eg, aggressive fluid and hemodynamic support, ventilatory support, and renal replacement therapy), use of investigational therapies, healthier patients, or chance. The lack of a control group, differences in the patients' disease severity, and use of multiple interventions (and varied courses of therapy) have made it impossible to draw definite conclusions. Further complicating matters have been uncertainties about the interplay between the virus and factors influencing its lethality. Even if the mortality rate had been more extreme (eg, 10% or 90%), the question of treatment efficacy would not have been answered definitively, raising questions about potential selection bias from inclusion of only minimally infected (in the case of low mortality) or of severely diseased (in the case of high mortality) cases.

A randomized controlled trial (RCT) is a proven and efficient way to learn whether experimental treatments offer any true benefit over supportive therapy. The Ebola Medical Countermeasures (MCM) study, structured to enroll patients in both the United States and West Africa, incorporates an adaptive, randomized, controlled design intended to address the safety and efficacy of the most promising EVD therapies.

The design of a clinical trial during this outbreak faced unique challenges including, but not limited to (1) severely limited and intermittent supply of some treatments due to such

Received 16 October 2015; accepted 4 February 2016; published online 11 February 2016.
Correspondence: L. E. Dodd, National Institute of Allergy and Infectious Diseases, 6700 B Rockledge, 5601 Fishers Lane, MSC 9820, Bethesda, MD 20892 (doddl@mail.nih.gov).

The Journal of Infectious Diseases® 2016;213:1906–13
 DOI: 10.1093/infdis/jiw061

factors as the complexity of the manufacturing process (eg, requiring producers to ramp up production to meet the projected needs of study enrollments) or the genetic evolution of the virus (eg, requiring producers to better match anti-sense compounds to the prevalent Makona strain); (2) considerable uncertainty about the contemporaneous control mortality rate, which could be high but may vary depending on such factors as changes in best supportive care, regional variation in medical capacity, viral evolution, and timing of presentation at clinic from disease onset; (3) uncertainty about the number of potential participants in any given region, depending on epidemiologic and logistical factors; and (4) the potential need to stop earlier than might be typical for an RCT, if the experimental therapy looks sufficiently promising.

Rather than precluding an RCT, these circumstances should actually encourage it, to ensure valid conclusions about the benefits of experimental therapies. Nonetheless, an unusual amount of flexibility in trial design is needed to seamlessly accommodate changing circumstances. Although flexibility is required in a disease setting with tremendous dynamic pressure, limiting trial adaptations is important to ensure that reasonable conclusions can be reached. The design of the EVD MCM trial is described herein, followed by computer simulation studies that evaluate the properties of the design under various experimental settings.

METHODS

Study Design

Table 1 describes the basic features of this design. During the planning stages of this trial, investigational therapies were ranked according to the preexisting evidence that, at the time, was based largely on in vitro testing, small-animal model studies, and/or limited nonhuman primate studies. Based on this data, one candidate, ZMapp (Mapp Biopharmaceuticals) seemed to distinguish itself from the others [5]. For this reason, the current trial commenced with randomization to ZMapp plus supportive care versus best supportive care alone. Had several agents appeared equally promising, the design could have included multiple experimental arms and a best supportive care arm and would have used multiarm-multistage methods [6, 7]. Of note, the inclusion of ZMapp as the first candidate therapy to be tested in this RCT involving infected patients actually preceded phase 1 safety and pharmacokinetic testing of this product in normal human volunteers. Randomization was stratified by 2 factors—cycle threshold and region—to ensure balance across these potential prognostic factors. Secondary analyses will consider other prognostic variables, such as age.

Table 1. EVD Trial Characteristics

Study Endpoint	28-day Mortality
Treatments	Experimental: cycle through ranked list of candidate therapies or allow concurrent evaluation if therapies are equally ranked; control: concurrent control group of oSOC; control group may change if evidence accumulates to support superiority of a given investigational product and therefore incorporation of that element into new background oSOC
Randomization	Equal allocation to each study arm; permuted blocks of small size
Stratification factors	Location (Liberia or Sierra Leone, Guinea, United States); disease severity (cycle threshold, >22 of ≤22)
Monitoring	Aggressive early monitoring, with the possibility of stopping the trial, under extreme circumstances after enrollment of only 6 subjects per arm
Type I error rate	Less strict than usual
No. of patients	Target of up to 100 per arm, recognizing that final analysis will occur if epidemic ends before complete enrollment

Abbreviations: EVD, Ebola virus disease; oSOC, optimal standard of care.

The design implications of the features listed in the introduction can be seen in Figure 1. The figure shows that although the initial randomization is to optimized standard of care (oSOC, defined as the standard of care that reflects best regional practices in supportive care, with a minimum requirement of intravenous fluids, hemodynamic and electrolyte monitoring, and adjunctive medications such as antimalarials and antibiotics) or drug X + oSOC, a protracted shortage in supply of drug X may lead to alternative randomization to oSOC or drug Y + oSOC while the supply of drug X is being replenished.

Because the control mortality rate is unknown and could change over time, it is critical that the comparison of drug X + oSOC with oSOC should include only patients concurrently randomized to oSOC; patients randomized to oSOC during the "white bar" period of Figure 1 would not be included for that comparison but would be included only for the comparison between drug Y + oSOC and oSOC. Stratification by region (eg, United States vs West Africa) was implemented because the oSOC may vary greatly by location. Frequent interim monitoring would allow the trial to stop if early treatment outcomes strikingly favor the experimental arm, as described in Results. If that happens, the experimental agent would probably become part of the updated oSOC (subject to supply constraints) to be compared with updated oSOC + new experimental agent(s). At that point, the sample size may need recalculation to reflect a smaller expected intervention effect if the revised oSOC then includes a known effective treatment.

"Barely Bayesian" Trial Design

A Bayesian approach was taken to accommodate the need for flexibility. The Bayesian approach is different from the standard frequentist approach (ie, the paradigm that emphasizes *P* values for evidence), in that the Bayesian approach incorporates a prior belief about efficacy of treatments. These prior beliefs affect the estimates of efficacy with the use of a "prior" distribution, which is updated to a "posterior" distribution after observing data. Much controversy between these frequentist and Bayesian approaches has existed over the years largely owing to debate about the appropriate use of (or lack thereof) a prior distribution.

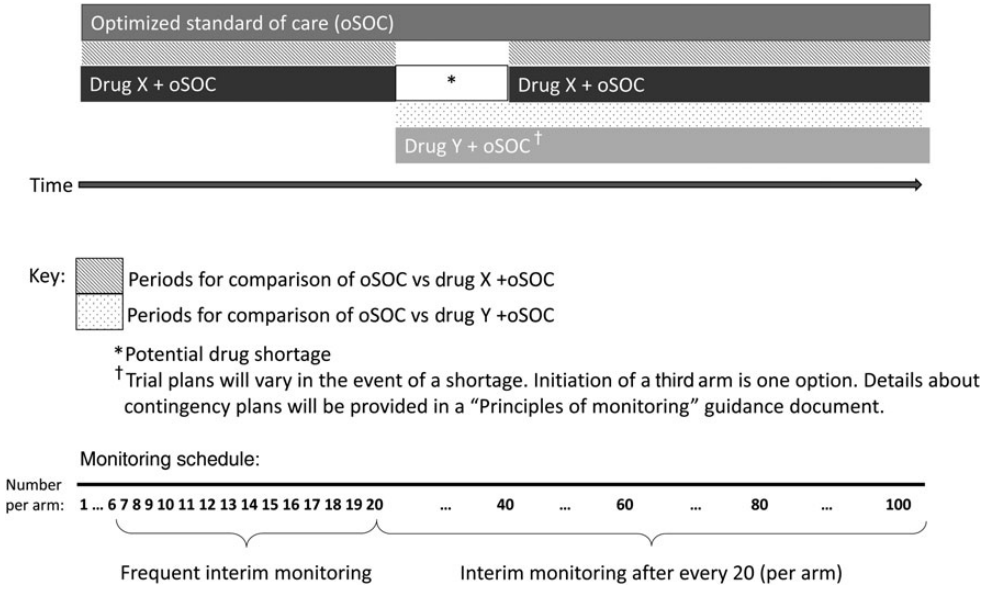

Figure 1. General study schema.

Thoughtful specification of the prior distribution is crucial in Bayesian analysis, because conclusions should depend primarily on data from the trial, not on prior opinion. A noninformative prior distribution that is quickly overshadowed by actual data was chosen, hence the term "barely Bayesian design." Furthermore, for a given prior, a design can be evaluated from a frequentist perspective by evaluating type I error rate (ie, the probability of wrongly concluding efficacy when the treatment has no benefit) and power (ie, the probability of correctly concluding efficacy when the treatment has benefit).

For the MCM design, a prior distribution on the probability of survival in a given arm was formulated as follows. Imagine having data on 2 persons treated with a given agent, and observing that 1 of the 2 survived. The probabilistic equivalent is to assume a uniform distribution on the interval (0, 1), which is consistent with an overall survival probability of 0.50 for the current Ebola outbreak but with wide variability reflecting substantial uncertainty. Moreover, a uniform distribution ensures very little influence of prior opinion on conclusions. Indeed, the observed data very quickly dominate in decision making. Figure 2 displays the Bayesian thought process. Figure 2A shows the initial distribution. Figure 2B and 2C show 2 updates to the distribution based on hypothetical data from 10 and 40 observations, respectively. The distribution in Figure 2C is much more concentrated around a 20% mortality rate because that posterior is based on more data than that in Figure 2B.

The primary analysis is based on calculating the posterior probability that the 28-day survival rate is higher in the investigational arm (arm X) than in the control arm (oSOC) (see Supplementary Material for formula). If the posterior probability is high at the end of the trial (eg, >97.5%), a conclusion of efficacy can be made. However, given the exigent circumstances of the EVD outbreak, it was believed that frequent and early monitoring would be needed. Therefore, posterior estimation is undertaken as the trial is ongoing, at specified time points. At any interim analysis preceding the final analysis, arm X is declared superior if there is ≥99.9% probability that its survival rate exceeds that of the oSOC arm. At the final analysis, superiority of arm X is declared if this probability exceeds 97.5%. Table 2 shows stopping criteria for sample sizes of 6–10 per arm. For example, after 8 patients have been enrolled in each arm, the following scenarios would cross the stopping boundaries: (1) 0 deaths in arm X and either 7 or 8 deaths in the oSOC arm or (2) 1 death in arm X and 8 for oSOC. (Supplementary Figure 1 in the Supplementary Material provides stopping boundaries for a broader set of scenarios.)

One potential concern that arises from basing decision making on posterior probabilities instead of the usual type I error rate is the probability of declaring a treatment beneficial when it is not. It would be disconcerting if a completely ineffective treatment had a high probability of being declared superior to control. This possibility is avoided by stopping for benefit at an interim analysis only if the posterior probability is quite high, 99.9%. Results from computer simulation studies allow evaluation of the type I error rate and are presented in "Results" section.

Futility Monitoring

Although we monitor frequently in hope of quickly identifying a successful intervention, the intervention may be poor or even harmful. There is a substantial risk inherent in randomizing subjects to an ineffective treatment when other, potentially promising, treatments are also available and, as a result, there

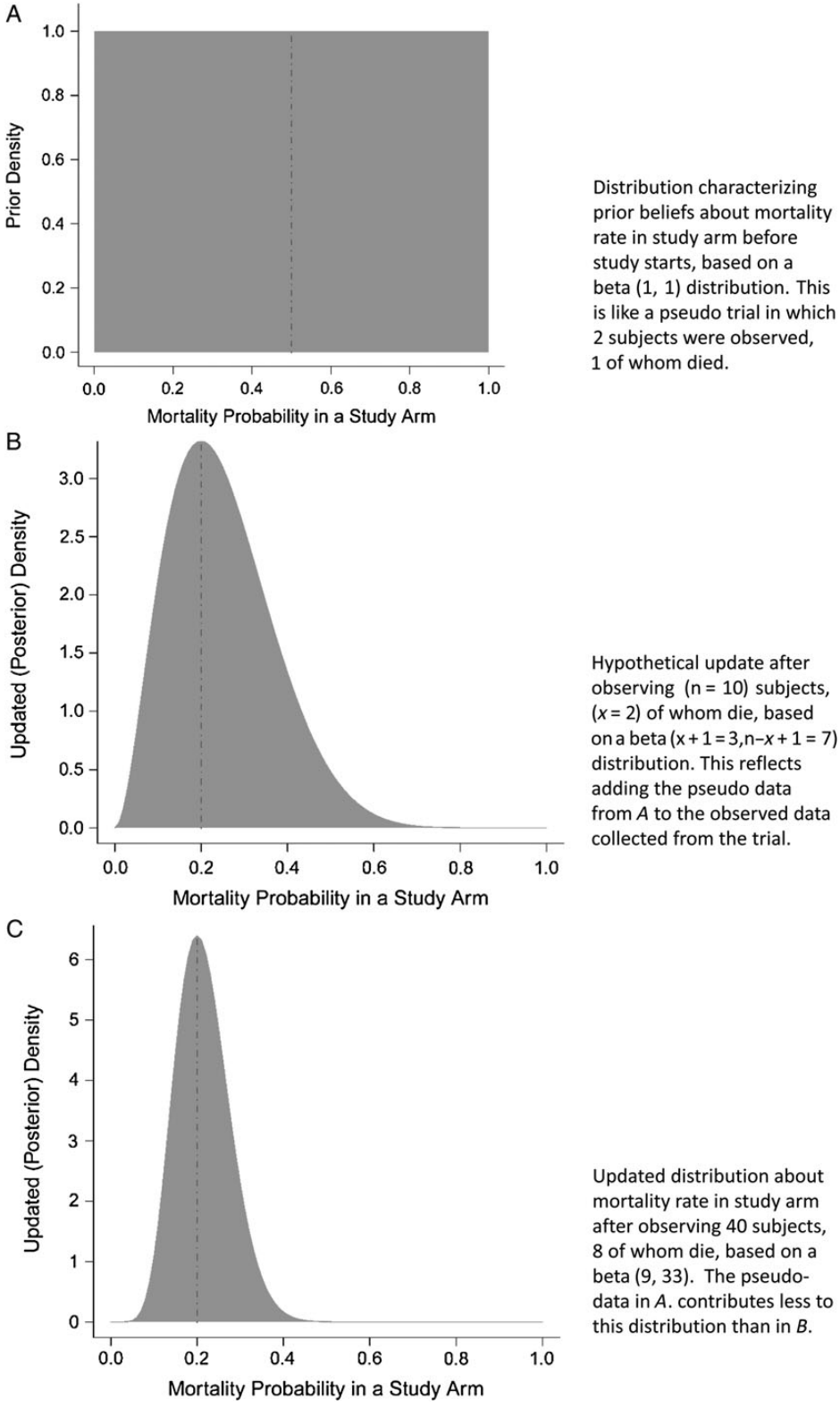

Figure 2. Bayesian thought process.

is an ethical imperative to terminate the trial early in the presence of overwhelming evidence for futility. We use conditional power to monitor for an ineffective intervention. Conditional power is the probability of achieving statistical significance at the end of the trial, given the data observed thus far. It can be computed under different assumptions about the treatment effect for future

Table 2. Stopping Boundaries by Number of Subjects per Arm for 6–10 Subjects per Arm

Stopping Boundary, No. of Observed Deaths									
6 Subjects per Arm		7 Subjects per Arm		8 Subjects per Arm		9 Subjects per Arm		10 Subjects per Arm	
oSOC	Arm X	oSOC	Arm X	oSOC	Arm X	oSOC	Arm X	oSOC	Arm X
6	0	6	0	7	0	7	0	7	0
. . .	. . .	7	(1 or 0)	8	(1 or 0)	8	(1 or 0)	8	(1 or 0)
. . .	. . .	. . .	. . .	. . .	. . .	9	(2 ,1, or 0)	9	(2, 1, or 0)
. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	10	(3, 2, 1, or 0)

Abbreviation: oSOC, optimal standard of care.

data. Most relevant for futility monitoring is to assume the treatment effect specified in the protocol, namely a 50% mortality reduction. The effect of conditional power on the properties of the design was also evaluated via simulation studies.

RESULTS

Figure 3 provides a timeline for this study, starting with a series of meetings on potential medical countermeasures that date back to September 2012. Protocol development began in September 2014 and proceeded with expedited reviews, taking about 7 months from initial protocol team discussions to study initiation. This is in contrast to a more typical timeline, which may take years [8, p 128; 9, chap 3; 10]. Sufficient supply of study drug was available in February of 2015. The first patient was enrolled in the United States on 13 March 2015, followed by a patient enrolled in Liberia on 22 March 2015. The World Health Organization reported no Ebola cases in Liberia the week from 23 March to 29 March 2015 and declared the outbreak over in that country on 9 May 2015, although the study remained open, which allowed enrollment of additional patients during small outbreaks in July and again in November. The first patient from Sierra Leone was enrolled in early April, followed by enrollments in Guinea, which started in early July. The study was intended to continue to enroll patients until the epidemic ends, full enrollment (of 200 subjects in total) was achieved, or a stopping boundary was met. In the event that the epidemic ended before achieving stopping criteria, the study was to be closed and a posterior probability of 97.5% was to be the boundary to declare efficacy.

An important feature of the barely Bayesian approach (BBA) is its ability to stop the study earlier under extreme circumstances relative to more conventional approaches. This is an important difference when compared with standard boundaries. Consider, for example, the O'Brien-Fleming boundary. If everyone survives in arm X and everyone dies under oSOC, the O'Brien-Fleming boundary [11] is not crossed until there are 15 in each arm. That is, if there are 14 per arm, 100% of whom survive in arm X and die under oSOC, the boundary is not crossed. By contrast, the BBA boundary is crossed with 6 per arm if all 6 survive in one arm and all 6 die in the other. The Haybittle-Peto boundary [12] in conjunction with Fisher exact test [13] is another alternative, but it is slightly more conservative, requiring 14 total subjects and 7 deaths in the oSOC and none in arm X before it can be crossed. Figure 4 describes the probability of rejecting under different settings of a treatment benefit for BBA relative to Fisher exact test with the Haybittle-Peto boundary. The biggest gain from BBA occurs if the treatment benefit is large, when stopping for benefit may occur as with as few as 6 per group. The largest difference can be seen for an extreme case of 80% mortality with the oSOC and a reduction to 20% mortality with arm X. After 7 subjects per arm, the probability of stopping is 20% with the BBA approach but only 5% using the Fisher exact test in conjunction with the Haybittle-Peto method. Supplementary Figure 2 in the Supplementary Appendix provides a similar graph but for the type I error rate.

Computer simulations with randomly generated outcomes were used to evaluate the BBA under various mortality rates. Table 3 summarizes the type I error rate (ie, the false-positive rejection rate) and power under various baseline mortality rates, with and without treatment effects. Note that futility monitoring, as described in Methods, was to be implemented after enrollment of 40 subjects per arm. One reasonable initial estimate of the mortality rate in the control arm was 0.50. If that turns out to be accurate, and if ZMapp has no effect, the false-positive probability is .033, higher than the conventional level of 0.025. This was considered acceptable under the circumstances. The error rates are somewhat lower if the control mortality rate differs from 0.50. The average sample sizes range from 50 to 90, and are all <100, owing to early stopping from low conditional power.

If the treatment is effective and the mortality rate in the control arm is about twice that in the ZMapp arm, power is reasonable. For example, a scenario with a mortality rate of 0.4 in the control arm and 0.2 in the arm with ZMapp was considered most reasonable when designing the study. Under this scenario, power is approximately 87% with an average sample size of 80 subjects per arm. Under a more extreme scenario of 80% mortality in the control group and 40% with ZMapp, power approaches 100%, with an average sample size of 36 per arm.

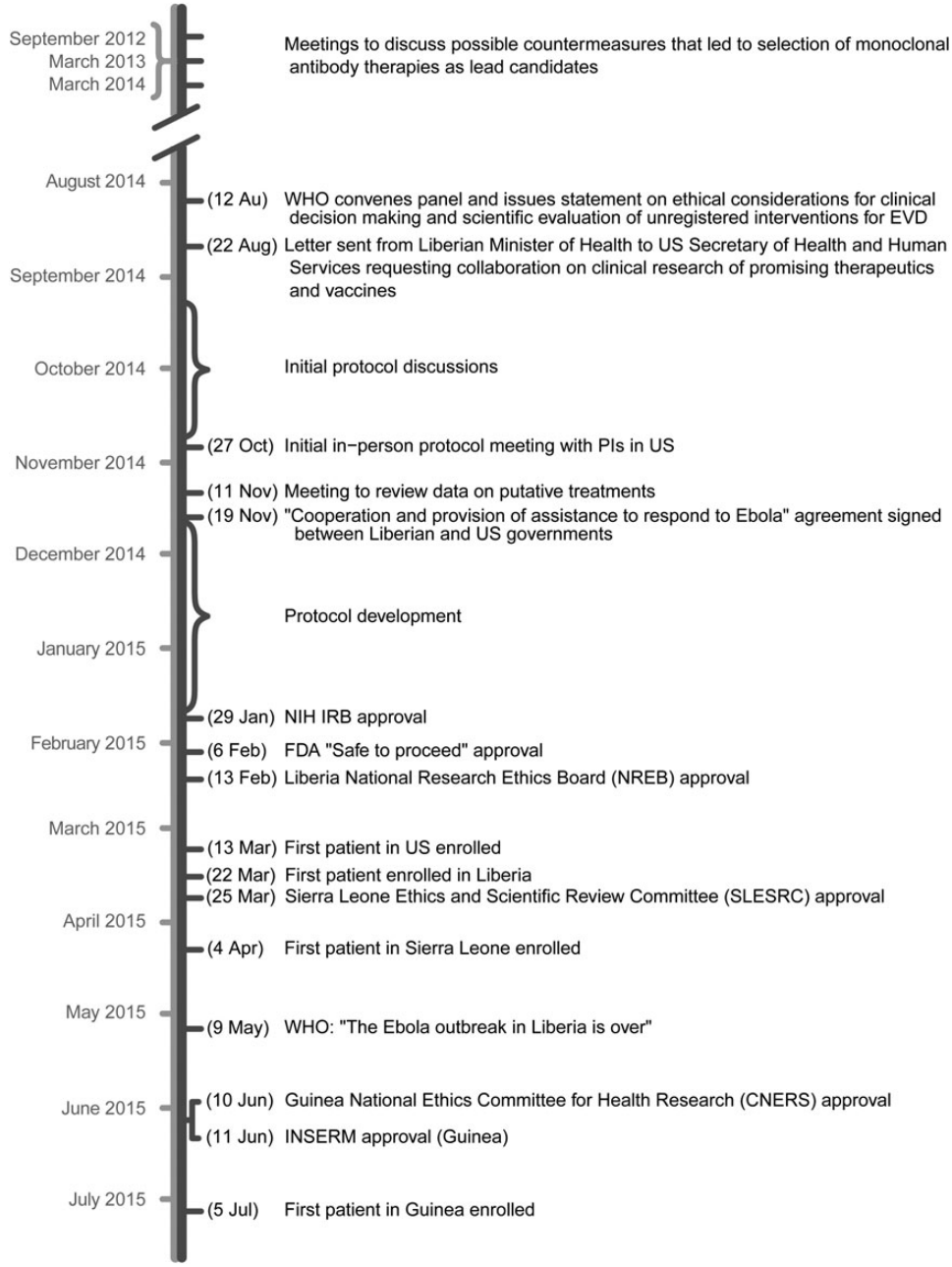

Figure 3. Study timeline. Abbreviations: EVD, Ebola virus disease; FDA, Food and Drug Administration; IRB, institutional review board; NIH, National Institutes of Health; PIs, principal investigators; WHO, World Health Organization.

Note that the study was designed for a total sample size of 100 per group. In different settings, the methods should be adapted to other settings based on the targeted effect sizes.

DISCUSSION

Clinical studies commonly require a compromise between the ideal scientific experiment and what is both logistically practical and ethically acceptable. The lack of an effective EVD treatment, limited understanding of the factors influencing mortality rates in different clinical settings, and existence of only anecdotal reports of outcomes from emergency investigational new drug use led to our decision to design and conduct an RCT. The approach used in the MCM study represents a compromise that includes randomization but allows early stopping if an experimental treatment is found to be exceedingly effective.

In a trial with an expected high mortality rate, it is imperative to stop for benefit as early as possible to make the treatment immediately available. In this study, an observation of 6 deaths

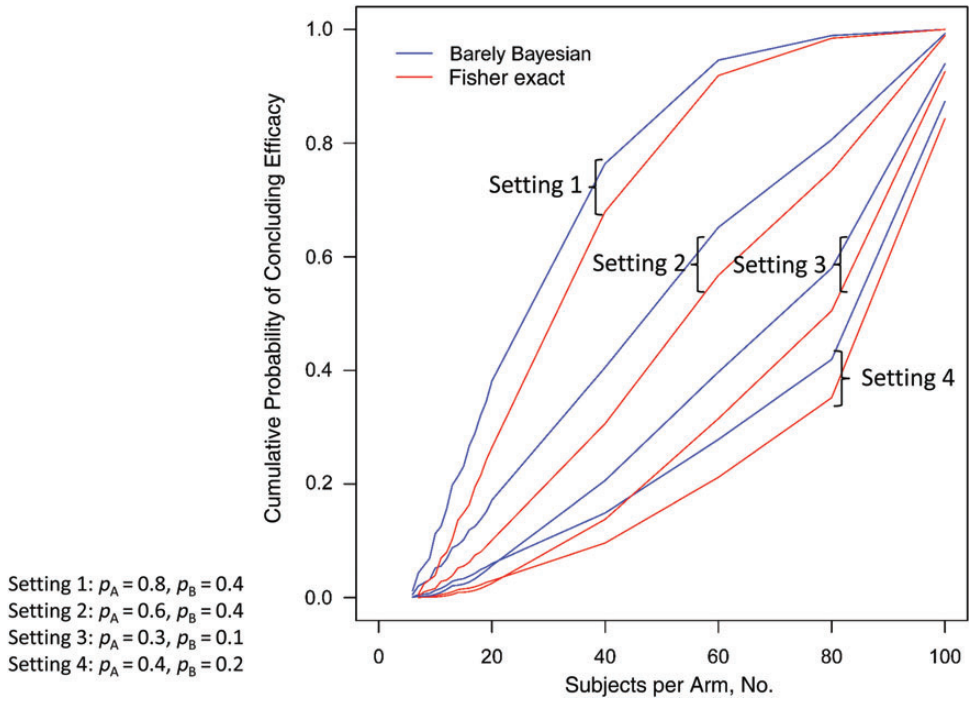

Figure 4. Power according to number of subjects per arm for 4 settings. p_A is the mortality rate in arm A. p_B is the mortality rate in arm B.

Table 3. False-Positive (Type 1) Error Rate and Power of Proposed Procedure for Different True Mortality Probabilities in the 2 Arms

Mortality Rate, oSOC	Mortality Rate, Arm X	Rejection Rate	Mean Sample Size per Group	Proportion of Times Group Sample Size Is		
				≤40	≤60	≤80
Under the Null Hypothesis (Type I Error Rate)						
0.1	0.1	0.018	49.7	0.69	0.88	0.94
0.2	0.2	0.023	58.9	0.41	0.73	0.91
0.3	0.3	0.027	67.0	0.19	0.58	0.87
0.4	0.4	0.029	73.9	0.07	0.42	0.81
0.5	0.5	0.033	79.9	0.02	0.24	0.74
0.6	0.6	0.028	85.2	<0.01	0.09	0.63
0.7	0.7	0.027	90.3	<0.01	0.02	0.46
0.8	0.8	0.025	88.9	<0.01	0.03	0.51
0.9	0.9	0.023	76.3	0.08	0.33	0.77
Under a Treatment Effect (Power)						
0.10	0.05	0.200	63.4	0.47	0.63	0.72
0.20	0.10	0.474	80.1	0.17	0.33	0.48
0.30	0.15	0.708	84.5	0.11	0.23	0.40
0.40	0.10	>0.999	48.9	0.55	0.82	0.94
0.40	0.20	0.873	80.1	0.15	0.29	0.47
0.40	0.28	0.432	87.8	0.05	0.16	0.37
0.50	0.25	0.958	70.9	0.25	0.44	0.62
0.50	0.30	0.834	82.8	0.13	0.25	0.40
0.50	0.35	0.590	89.6	0.06	0.14	0.29
0.60	0.30	0.993	58.2	0.41	0.65	0.81
0.80	0.40	>0.999	36.0	0.76	0.95	0.98
0.80	0.20	>0.999	15.5	0.998	0.999	1.000

Abbreviation: oSOC, optimal standard of care.

(among 6 subjects) in the oSOC arm vs no deaths (among 6 subjects) in the experimental arm was deemed extreme enough to stop the trial by investigators in this trial. We note that some might argue that more extreme results (such as 10 deaths among 10 subjects randomized to oSOC vs no deaths among 10 subjects randomized to experimental therapy) would have been more appropriate. Although consensus might not always be reached on this issue, the specific criteria embraced will likely vary depending on the circumstances and should be discussed extensively before trial initiation.

Additional features allowed adaptation of the control group and evaluation of multiple experimental agents simultaneously if multiple putative therapies were equally ranked. If >1 investigational product had ranked as leading candidates, randomization to multiple leading candidates versus supportive care would have featured prominently in the design. A strategy that evaluates multiple putative therapies would have efficiency gains, largely due to sharing of a common control group. However, this would require considerable coordination of efforts that may be difficult to achieve during an outbreak setting. When considering how many therapies can be reasonably evaluated, one must be cognizant of the inherent tension between evaluating a more interventions in fewer patients versus studying fewer interventions in more patients. A balance must be obtained to ensure definitive conclusions by limiting inclusion to only the most promising treatments.

Controversy over the ethics of an RCT in this context has been substantial [14, 15]. In addition to arguments based on scientific rigor, one argument in favor of this approach is that a lottery is the most ethical means of dispensing a product when, as in the early case with ZMapp, demand exceeded supply. This argument is reminiscent of the tuberculosis streptomycin study run by the British Medical Research Council in the 1940s, the original RCT that led to acceptance of the RCT as the reference standard for evidence-based medicine. Limited supply was one justification for the use of controls in that study.

Concerns that evaluations could be confounded by mortality rates changing over time, by site or by unknown factors, are further reasons for conducting an RCT. In this regard, Lindblade et al [16], who evaluated 9 remote outbreaks in Liberia, report a decline in mortality rates from 67% in August and September 2014 to 50% in December 2014. Further data from Kenema and Freetown in Sierra Leone revealed similar trends. In Kenema, the mortality rate was 74% from May to June 2014 [3], whereas in Freetown the rate was 48% from September to October 2014, with a decline to 23% mortality from November to December 2014 [4]. That said, the use of controls, particularly the most appropriate type of controls, in the context of an epidemic with high and potentially shifting mortality rates, will continue to be debated. The goal of rapid accumulation of reliable evidence about efficacy in order to make proven, effective treatments available as quickly as possible during an outbreak points to a unconventional trial with adaptive elements and reliance on the best known supportive care as a most appropriate initial comparator.

In January 2016, the current trial was closed owing to the decline in new EVD cases in West Africa. Thus, the full scope of this design will probably not be fully tested in the current outbreak. However, the conceptual flexibility incorporated herein may nonetheless be needed in new EVD studies during future outbreaks or in other emerging infectious diseases outbreaks as they continue to appear in parts of the world lacking effective prevention and control measures.

Supplementary Data

Supplementary materials are available at http://jid.oxfordjournals.org. Consisting of data provided by the author to benefit the reader, the posted materials are not copyedited and are the sole responsibility of the author, so questions or comments should be addressed to the author.

Notes

Acknowledgments. We acknowledge the extraordinary efforts of the PREVAIL II study team and input from the US Food and Drug Administration.

Disclaimer. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the US government.

Financial support. This work was partially funded by the Division of Clinical Research, National Institute of Allergy and Infectious Diseases. This project has been funded in part with federal funds from the National Cancer Institute, National Institutes of Health (contract HHSN261200800001E).

Potential conflicts of interest. All authors: No reported conflicts. All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

References

1. WHO consultation on potential Ebola therapies and vaccines. http://apps.who.int/iris/bitstream/10665/136103/1/WHO_EVD_Meet_EMP_14.1_eng.pdf?ua=1. Accessed 2 March 2016.
2. Ethical considerations for use of unregistered interventions for Ebola viral disease. http://apps.who.int/iris/bitstream/10665/130997/1/WHO_HIS_KER_GHE_14.1_eng.pdf?ua=1. Accessed 2 March 2016.
3. Schieffelin JS, Shaffer JG, Goba A, et al. Clinical illness and outcomes in patients with Ebola in Sierra Leone. N Engl J Med **2014**; 371:2092–100.
4. Ansumana R, Jacobsen KH, Sahr F, et al. Ebola in Freetown area, Sierra Leone—a case study of 581 patients. New Engl J Med **2014**; 372:587–8.
5. Jahrling PB, Hensley LE, Barrett K, Lane HC, Davey RT. State-of-the-art workshops on medical countermeasures potentially available for human use following accidental exposures to Ebola virus. J Infect Dis **2015**; 212(suppl 2):S84–90.
6. Royston P, Parmar MK, Qian W. Novel designs for multi-arm clinical trials with survival outcomes with an application in ovarian cancer. Stat in Med **2003**; 22:2239–56.
7. Proschan MA, Dodd LE. A modest proposal for dropping poor arms in clinical trials. Stat Med **2014**; 33:3241–52.
8. Nass SJ, Moses HL, Mendelsohn J, eds. Implementing a National Cancer Clinical Trials System for the 21st Century. Washington, D.C.: National Academies Press, **2013**. http://www.nap.edu/read/18362/chapter/1.
9. English RA, Lebowitz Y, Giffin RB. Transforming Clinical Research in the United States. Washington, D.C.: National Academies Press, **2010**. http://www.nap.edu/read/12900/chapter/1.
10. Rosas SR, Schouten JT, Dixon D, et al. Evaluating protocol lifecycle time intervals in HIV/AIDS clinical trials. Clin Trials **2014**; 11:553–9.
11. O'Brien PC, Fleming TR. A multiple testing procedure for clinical trials. Biometrics **1979**; 35:549–56.
12. Haybittle JL. Repeated assessment of results in clinical trials of cancer treatment. Br J Radiol **1971**; 44:793–7.
13. Agresti A. A survey of exact inference for contingency tables. Stat Sci **1992**; 7:131–53.
14. Cox E, Borio L, Temple B. Evaluating Ebola therapies—the case for RCTs. N Engl J Med **2014**; 371:2350–1.
15. Cooper BS, Boni MF, Pan-ngum W, et al. Evaluating clinical trial designs for investigational treatments of Ebola virus disease. PLoS Med **2015**; 12:e1001815.
16. Lindblade KA, Kateh F, Nagbe TK, et al. Decreased Ebola transmission after rapid response to outbreaks in remote areas, Liberia, 2014. Emerg Infect Dis **2015**; 21:1800–7.