

The Value of the Information That Can Be Generated: Optimizing Study Design to Enable the Study of Treatments Addressing an Unmet Need for Rare Pathogens

Aaron Dane,¹ John H. Rex,² Paul Newell,² and Nigel Stallard^{3,✉}

¹DaneStat Consulting, Cheshire, United Kingdom, ²F2G Limited, Manchester, United Kingdom, and ³Warwick Medical School, University of Warwick, Coventry, United Kingdom

In traditional phase 3 trials confirming safety and efficacy of new treatments relative to a comparator, a 1-sided type I error rate of 2.5% is traditionally used and typically leads to minimum sizes of 300–600 subjects per study. However, for rare pathogens, it may be necessary to work with data from as few as 50–100 subjects. For areas with a high unmet need, there is a balance between traditional type I error and power and enabling feasible studies. In such cases, an alternative 1-sided alpha level of 5% or 10% should be considered, and we review herein the implications of such approaches. Resolving this question requires engagement of patients, the medical community, regulatory agencies, and trial sponsors.

Keywords. alternate statistical criteria; antibiotics; antifungals; small trials; unmet need.

Traditional phase 3 trials confirm safety and efficacy of a new treatment relative to an active comparator or placebo using a 1-sided type I error rate of 2.5%. Although the 2.5% type I error rate is not a regulatory requirement, it is traditionally used by regulatory agencies to limit false conclusions of treatment effect. Given the high unmet need, in some infectious disease research the US Food and Drug Administration (FDA) have considered flexibility in benefit-risk considerations, granting approval with a single adequate and well controlled trial, with supportive information such as a phase 2 trial or information from in vitro studies and animal models of infection. Although this enables progression of some clinical development programs, there are still situations when it may only be feasible to recruit 50–100 patients in many years [1], for example, the plazomicin bacteremia study for carbapenem-resistant *Enterobacteriaceae* [2], making even a single trial extremely challenging. As a result, there are no, or very limited, data in such patient populations, so development of agents in small populations requires alternative approaches. More importantly, we only envision these approaches being used when a larger trial is simply not possible, so these ideas do not apply when there is a realistic choice between a larger or smaller study.

Drawing on value of information ideas within the orphan drug area [3–6], we propose an approach that uses alternative type I error rates in a decision-making and sample-size determination framework in which typically sized controlled studies are extremely challenging to enroll. The goal is to balance the risks of correct and incorrect decisions from limited data. Both approval and nonapproval carry risks, and we propose an approach that considers risks of failing to approve a new effective medicine, approval of an ineffective medicine, and conducting a large randomized controlled trial ([RCT] with some patients receiving suboptimal therapy) in a limited population.

We discuss regulatory approval on the basis of efficacy, because the aim is to understand the risks of using alternative statistical criteria within an efficacy assessment. The safety and benefit-risk profile are also critically important and need to be considered in any evaluation, but because the safety profile is unique to each treatment, we have not discussed this aspect in detail, but it is always a critical aspect of any evaluation of the evidence to support regulatory approval.

Superiority Trials Are Not Appropriate for All Development Programs

It is often perceived that demonstration of superiority in small trials with resistant pathogens should be straightforward if a strong effect exists. However, this is infeasible because we cannot study ineffective comparators in seriously ill patients, and demonstration of superiority is challenging when therapies exist with some efficacy, albeit suboptimal [7, 8]. Although combinations of less effective therapies may have toxicities, their existence makes superiority hard to demonstrate.

Received 02 March 2022; editorial decision 19 May 2022; accepted 25 May 2022; published online 27 May 2022

Correspondence: A. Dane, MSc, DaneStat Consulting, 161 Park Lane, Macclesfield, Cheshire, UK (aaron.dane@danestat.com).

Open Forum Infectious Diseases®

© The Author(s) 2022. Published by Oxford University Press on behalf of the Infectious Diseases Society of America. This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs licence (<https://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial reproduction and distribution of the work, in any medium, provided the original work is not altered or transformed in any way, and that the work is properly cited. For commercial re-use, please contact journals.permissions@oup.com
<https://doi.org/10.1093/ofid/ofac266>

In some settings, the use of a nonstandard significance level to assess superiority may make studies more feasible, but in many cases the issues associated with demonstrating superiority described above will still exist, meaning alternative options are also required. Therefore, we focus on noninferiority (NI) trials, but similar arguments to those presented could be used to assess the impact of changes to type I error in superiority trials.

Decision Making and Sample Size in Randomized Controlled Trials With Rare Pathogens

The challenge reduces to recognizing that rare pathogens make it impossible for practical purposes to conduct either a small superiority study or a large noninferiority active-control study. However, new therapies are needed for rare pathogens, and we can approach this problem by being sure that we are confident of the following: (1) we will approve effective treatments (showing superiority or noninferiority with an active control); and (2) we will not approve ineffective treatments.

In registrational studies, it is traditional to set power (probability of correctly showing the new agent is effective) at 80%–90% and type I error at 2.5% (1-sided, the probability of concluding a new treatment is effective when it is not) and set sample size accordingly [9]. Table 1 summarizes these concepts and the consequences for patients within the trial and those beyond the trial who will receive the current or new therapy dependent upon the trial results.

When we design the trial, we do not know the true response rate, so we must understand the probability of approval for different study sizes and scenarios regarding true treatment effects. We are attempting to identify a “sweet spot” for sample size balancing the type I error and power when it is infeasible to obtain the traditional chances of correctly or incorrectly approving a new treatment.

Where possible, trials with conventional power and type I error are preferable. However, if the trial is too large, this deprives patients of timely access to an effective therapy and means that while the trial is ongoing, some patients within the trial will receive an inferior treatment. If a trial is too large, developing the therapy is infeasible because we cannot generate data for that patient population. In contrast, if a trial is too small, it increases the probability of making an incorrect decision (either approval or nonapproval).

We propose that when large trials are infeasible, alternative type I error rates should be explicitly considered within the context of the benefits and risks of new treatments, particularly when accompanied by supportive information such as data from in vitro studies and animal models of infection. We consider this for rare pathogens in instances in which new therapies are needed and patient numbers make clinical trials prohibitively lengthy or simply impossible to undertake. This has the potential to accelerate availability of new effective

Table 1. Scenarios to Consider When Analyzing Decision Making With Small Datasets

True Response Rate or Mortality Rate	Consequences Within RCT	Consequences for Patients After RCT
New agent is worse than control	Every patient randomized to new agent risks a worse outcome	If the new agent is approved, this problem is perpetuated, and all future patients would receive a less effective therapy <i>“We should avoid incorrect approvals so want to have low power of approval”</i>
New agent is better than control	Every patient randomized to control risks a worse outcome	If the new agent is not approved, this problem is perpetuated, and all future patients would receive a less effective therapy <i>“We should retain high power of approval”</i>
New agent is similar to control	Every patient has approximately the same outcome	Either existing treatment or a new therapy would be acceptable, but we still need to have additional therapies available. <i>“Therefore, we still want to retain high power of approval”</i>

Abbreviations: RCT, randomized controlled trial.

treatments rather than condemn patients to suboptimal treatment while larger trials are either attempted in vain or never attempted.

Understanding Clinical Trial Risks

In considering alternative statistical criteria, we first state key risks and then consider acceptable risk levels in an area of extreme unmet need. Risks must be considered for a range of scenarios in which the true effect of the test agent related to control is unknown. This is true of all studies but may need to be considered differently when trials cannot be designed to standard levels of power and type I error. Important scenarios are situations in which the new agent is better than, worse than, or similar to control. We consider these in turn for a randomized clinical trial using a 1:1 randomization to new agent or control with a 20% NI margin and 95% confidence interval [CI] (2-sided $\alpha = 0.05$), as has been used for some trials [10, 11].

Scenario 1: New Agent Better Than Control

Imagine the new agent and control have true response rates (unknown before the trial is conducted) of 60% and 40%, respectively. The correct outcome would be to conclude NI and approve the new agent. The probability that NI is concluded (power) with this scenario increases with sample size (N) (Figure 1A) but does not improve greatly with more than 40–50 patients/group, when power approaches 90%. Although more data may be needed from a safety perspective, this would

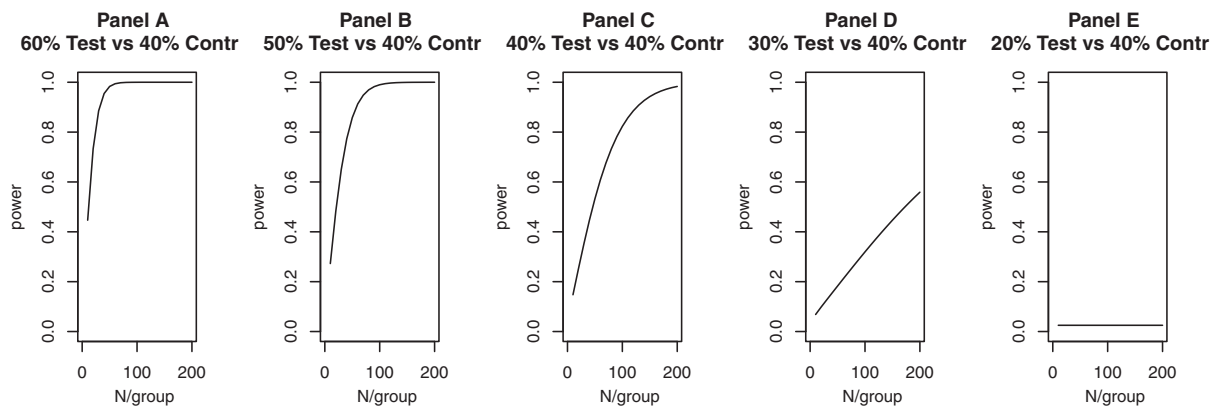


Figure 1. Plot of power when using 95% confidence interval and -20% noninferiority margin. Cont, control.

be sufficient to show the new agent has efficacy noninferior to control.

Scenario 2: New Agent Slightly Better Than Control

When the new agent and control have true response rates of 50% and 40%, respectively, the correct outcome would again be to conclude NI and approve the new agent. In this case, as presented in Figure 1B, the probability NI is concluded increases with sample size, but power is lower for a given sample size than for Scenario 1 and approaches 80% at 50 patients/group.

Scenario 3: New Agent Similar to Control

In the third scenario, the true (but as yet unknown) response rates are the same (40%) for new agent and control. The correct decision in this scenario is more nuanced, and we must consider the value of currently available therapies, even if these are suboptimal. As stated, we assume that it would be desirable to have additional therapies available.

In this case, the probability of concluding NI (power) increases with sample size (Figure 1C) but requires a larger sample size. A study with ~ 50 patients/group has $\sim 50\%$ power, whereas a study of 80 patients/group provides $\sim 80\%$ power. It would take a large study to achieve 90% power. Because it is unlikely that studies would be undertaken with only a 50/50 chance of success, developing new agents is not possible when it is infeasible to recruit more than 100 patients. If new therapies are badly needed, a discussion on required levels of type I error and power should be considered.

Scenario 4: New Agent Slightly Worse Than Control

When the true response rates for the new agent and control are 30% and 40%, respectively, the correct outcome and associated approval is less obvious. If the primary endpoint of the study relates to all-cause mortality, you would not want to approve such an agent, but with a response rate endpoint, if there is a high unmet need, the available treatment has significant toxicity, and

there is increasing resistance to this available treatment there could still be value in a new agent. You would not always want to approve such an agent, but consideration of these components of the available therapy would be important when considering the need for new agents. Therefore, with this scenario, there is no correct outcome for every situation. Rather, consideration of whether the new agent should or should not be approved will depend upon the reason why a new agent is needed. In this case, the plot shows the probability of concluding NI is $< 25\%$ for sizes of $< \sim 75$ /arm (Figure 1D). The probability of concluding NI increases as the sample size further increases, and at 200/arm the probability approaches 50%. However, as such sample sizes are presumed impossible for rare pathogens this again shows a relatively small chance of concluding NI with achievable sample sizes of 40–50 patients/group.

Scenario 5: New Agent Worse Than Control

When the true (but as yet unknown) response rates for the new agent and control are 20% and 40%, respectively, we assume we would definitely not want to approve the new agent, so the correct outcome is that the new agent should not be approved. In this case, the probability of concluding NI is always very low (Figure 1E). Because the true difference between treatments equals the NI margin, this is the type I error rate (2.5%), so we rarely conclude NI erroneously, whatever the study size. Therefore, 40–50 patients/group again seems reasonable.

The Effect on Future Patients With This Rare Condition

We have shown that trials studying infections caused by rare pathogens are infeasible with traditional power and type I error rates. Another important component is the implication of any drug approval decision on the treatments received by patients with this rare pathogen, both in the RCT and in the future.

As a hypothetical example, suppose we have an overall population of 1000 patients who might benefit from a new

treatment, and from this population we recruit 100 patients (50/arm) to a trial comparing the new agent with control, the results of which will define the therapy used for the remaining 900 patients, dependent upon the NI conclusion made from the trial (see Table 2).

Assuming all patients are included in the trial or are treated based on the results of the trial, we can calculate how many patients are successfully treated using the following. (1) How many of 50 trial patients on the new agent are successfully treated? (2) How many of 50 trial patients on Control are successfully treated? (3) How many of remaining 900 patients are successfully treated depending upon the trial outcome?

The expected number of successes can be calculated under the scenarios considered above. When the new agent response rate is 60%, we would expect 30 of 50 trial patients receiving the new agent and 20 of 50 trial patients receiving control to respond (Table 2). We would expect 540 of the remaining patients to respond if NI is concluded and the new agent is used and 360 of the 900 patients to respond on control if NI were not concluded.

Using a 20% NI margin and 95% CI, with true response rates of 60% and 40%, a trial with 50 patients/arm would have a 98.3% chance of concluding NI. Thus, we expect the new agent to be approved 98.3% of the time (so 540 of 900 future patients would respond) and the control to be used (so 360 of 900 respond) 1.7% of the time. The total expected number of responses from the 1000 patients is therefore 587 as displayed in the first row of Table 2. The second row of Table 2 shows similar calculations using the same assumptions but a larger RCT including 200 patients/group, leaving 600 patients treated based on the RCT results. In this case, the total expected number of responses is lower at 560 patients. This may seem counterintuitive, but although larger RCTs increase the probability of reaching the correct conclusion, they also mean we wait longer to reach this conclusion, meaning more patients within the trial are randomized to receive a less

effective therapy, and a smaller number of patients benefit from the more effective therapy after the trial.

The example in Table 2 is for the most optimistic scenario considered and presents results for trials with 50 and 200 patients per arm only. Similar calculations are possible for each of the 5 scenarios described and for different sample sizes and show a similar pattern of declining response with larger sample sizes for less optimistic scenarios (Figure 2). Recall that Figure 1 shows the probability of concluding NI in each case, whereas Figure 2 shows the proportion of responses expected in the population for each scenario. These results can be used to assess whether the risks of correctly approving an effective therapy, incorrectly approving an ineffective therapy, and the number of patients receiving suboptimal treatment have been appropriately balanced for a given trial size.

If the new agent is better than the control, a trial with 40 patients/group gives reasonable power and optimizes the expected successes in the overall population based on the treatment patients receive during and after the trial (Figure 2A). When the new agent is slightly better than control, the power is lower than when the new agent is clearly better, but it still declines after 40–50 patients/arm.

If the new agent is worse than control, as noted above, the probability of incorrectly recommending the new agent is low irrespective of the RCT sample size. The expected proportion of responses decreases with increasing sample size because more patients in the trial receive the (less effective) new agent (Figure 2E). When the new agent is slightly worse than control, a similar pattern is seen with the expected proportion of successfully treated patients, whereas the probability of concluding NI increases with increasing sample size, but there is still only 15%–20% probability of concluding NI with 40–50 patients/arm.

If the new agent and control have similar efficacy, the probability of recommending the new agent increases with N, and a larger study is required using traditional criteria. In this scenario, because the new agent and control have the same response rate, the expected proportion of patients to respond is constant (here 40%) (Figure 2C). Although the unmet need may appear less urgent when the efficacy of the new agent is similar to control, there are still a number of cases in which additional treatment options are urgently needed due to emerging resistance or toxicity concerns with current therapies. In this scenario, a study including ~100 patients/group would be required to demonstrate efficacy. In studies of some rare or resistant pathogens, even 40–50 patients/group may be challenging, so it is necessary to find a way of running a feasible trial. We next suggest a framework for assessing alternative statistical criteria building on ideas presented above.

A Framework for Using Alternative Statistical Criteria

We have shown that although conventional error rates are achieved using 40–50 patients/arm when the new treatment is truly clearly better or worse than the control, 100 patients/

Table 2. Expected Number of Responses in the Entire Population With an NI Trial of 100 or 400 Patients: Response Rates of 60% for the New Agent vs 40% for Control

RCT Sample Size	Responses in RCT		Probability New Agent Is Approved	Responses After Trial	
	New Agent	Control		If New Agent Is Approved	If New Agent Not Approved (Control Is Used)
50/arm	30/50	20/50	0.983	540/900	360/900
Expected Successes in Overall Population: $30 + 20 + (0.983 \times 540) + (0.017 \times 360) = 587$ (58.7% of all patients)					
200/arm	120/200	80/200	0.999	360/600	240/600
Expected Successes in Overall Population: $120 + 80 + (0.999 \times 360) + (0.001 \times 240) = 560$ (ie, 56% of all patients)					

Abbreviations: NI, noninferiority; RCT, randomized controlled trial.

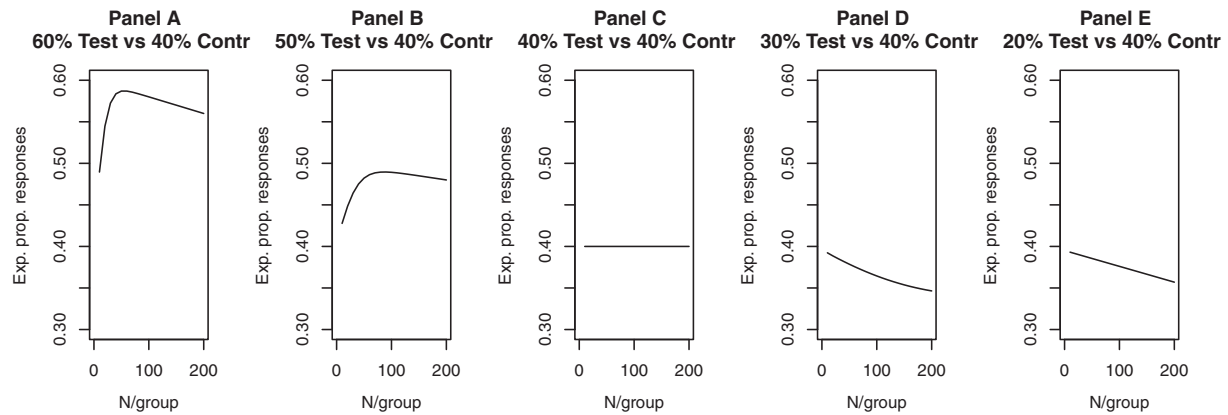


Figure 2. Plot of expected (Exp.) proportion (prop.) of responses in a trial using 95% confidence interval and -20% noninferiority margin. Cont, control.

arm may be necessary when the 2 treatments have similar efficacy. When such a study is infeasible, and additional treatment options are desirable, use of an alternative type I error rate (or CI) can be considered. Here we present the use of an 80% CI (2-sided $\alpha = 0.2$) and a 20% NI margin.

This approach would be reserved for circumstances with high unmet need and limited population where it is critical that new therapies are available but could not otherwise be developed. We will not discuss how an agent would meet these criteria, but we do outline how the risks associated with failure to make new therapies available, and the risk of incorrectly approving new agents, could be considered for such agents.

Properties of trials with an 80% CI are shown in Figure 3, which includes in one plot the type of information presented in Figures 1 and 2. When the new agent is better than control, the expected number of patients responding declines after ~ 50 patients/arm. The probability of a successful trial when the treatments are similar is higher (Figure 3, Panel C), approaching 80% with 50–60 patients/arm. The probability of such a trial incorrectly concluding NI when the new agent is worse (Figure 3, Panel E) is 10%, so it is above the usual value of 2.5% when using a 95% CI. This should be a point for discussion among patients, Infectious Diseases physicians, regulatory agencies, and sponsors for rare pathogens or for limited populations when additional treatment options are desperately needed. When the new agent is slightly worse, a similar pattern is seen to that when using a 95% CI in that the probability of concluding NI increases with increasing sample size, in this case meaning the chance of concluding NI approaches 40% with 50 patients/arm. As stated previously, the utility of an agent with the possibility of a slightly lower response rate but the ability to provide other benefits in terms of less toxicity or pathogen resistance should form part of the assessment of using an 80% CI.

The previous sections showed that when the new agent is clearly superior to the control, this would be apparent from

50 patients/arm. The alternative type I error rate level is important when the new agent and control are similar or slightly inferior. If it is only feasible to run a study of 100 patients, a trial with traditional type I error rate is impossible. Additional flexibility is then required, particularly when available therapies are suboptimal, such as being the drug of last resort or with toxicity issues where it is desirable to have additional therapies available. In this case, it will often be a choice between a smaller dataset or having no data at all.

DISCUSSION

Although studies of rare pathogens are challenging, the development of agents for such pathogens remains important. We have shown that although studies with traditional statistical criteria may be impossible, we can articulate a decision-making framework to support smaller studies. The proposed framework displays the risks of incorrect approval of an ineffective agent and incorrect rejection of an effective agent based on efficacy, along with the expected number of successes within a small population, highlighting that agreement needs to be reached across the community on how to balance these factors. The approach is similar to that proposed for orphan drug trials [3–6] in which a type I error rate larger than 0.025 is sometimes used [12, 13].

We have shown that comparative trials with 50 patients/arm can provide reliable decisions and could be the key to enabling any development for patients with rare pathogens. In our example, we have assumed that the control group event rate is 40%, and although our exact findings regarding the probability of approval may not hold if the control group rate were very different, similar methods could be used to choose an appropriate sample size.

We have also shown that a larger study does not always provide better outcomes in a limited population because fewer

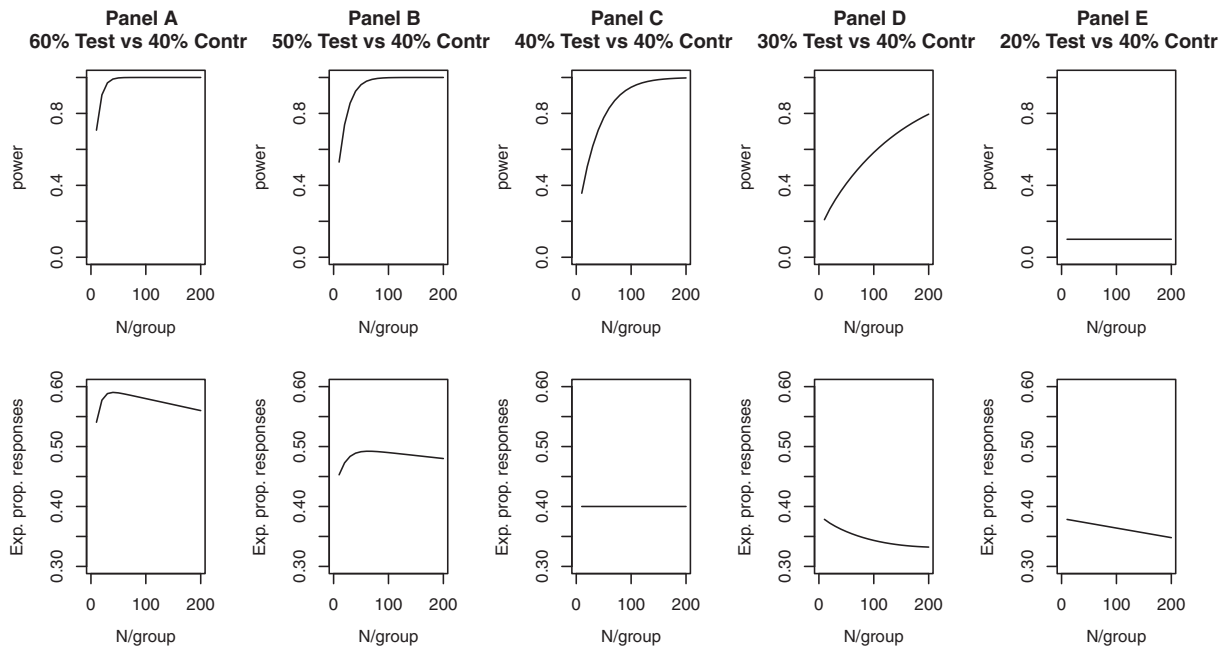


Figure 3. Plot of power and expected (Exp.) proportion (prop.) of responses in a trial using 80% confidence interval and -20% noninferiority margin. Cont, control.

patients remain to benefit from a therapy after the trial. In addition, in a trial with a 1:1 randomization in which one treatment is less effective, 50% of patients receive this suboptimal therapy.

We have presented an example in which an 80% CI is used alongside a 20% NI margin. Dependent upon the degree of unmet need, feasibility of recruitment, and the basis for justifying the NI margin, alternative approaches using different NI margins and levels of alpha (such as 90% CIs) could be considered to make trials more feasible. We acknowledge that the use of 80% or 90% CIs increases the risk of falsely concluding noninferiority when compared with the traditional 95% CI, and this is the reason this approach should be reserved for situations in which new treatments are urgently needed for a small patient population where a trial using a 95% CI is not possible. In these situations, additional flexibility may be warranted and would maintain scientific credibility by explicitly stating the error risk levels and would be particularly justified in the situation we envisage where the trial results are also augmented with robust information from in vivo and in vitro models of infection. We feel that this is also aligned with the intention stated in FDA guidance, where FDA can approve if there is “a positive benefit-risk balance in the limited population...even though insufficient data exist to conclude that there is a favorable benefit-risk profile in a broader population.”

Our calculation of patients benefitting from the chosen therapy assumed that all patients are either included in the trial or benefit from its results by receiving the preferred treatment. In reality, some patients will receive standard treatment outside

the trial while it is ongoing. Including these patients in the calculations of the expected proportion of patients responding would lead to even smaller trials being optimal when the new agent was effective (as even more patients are receiving a suboptimal therapy) and would make little difference when it was not effective. Therefore, as the trial sample size becomes large, the expected number of patients benefitting from the chosen therapy would decrease even more rapidly than in the calculations presented in this paper.

We acknowledge that this is a different level of evidence of efficacy than is usually used for regulatory approval, and this is why it is critical to gain the agreement of regulatory agencies and the infectious disease community when designing such studies that there is sufficient unmet and a small population which justifies this alternate approach to provide interpretable data for this patient population that would not otherwise be available. Furthermore, and as with any NI trial, it is important to be sure of the effectiveness of the control agent, and that the primary endpoint is objective and reflects a true clinical benefit to patients.

Beyond a change in statistical criteria for the primary endpoint, agreement to such a design and any final assessment would always consider the benefit-risk profile of a new agent. This would include detailed review of secondary efficacy endpoints. Knowledge of safety risks of an agent before the trial, or with a specific drug class, may impact the acceptability of a smaller safety dataset. These considerations are specific to a given agent but would always need to be considered when designing a trial as we have proposed.

This approach is relevant for drugs with activity only against a specific pathogen(s), and also for drugs with broad-spectrum activity which includes an uncommon pathogen. In both cases, this would provide important, quantitative data on pathogens of interest, with the key difference being the amount of additional information it is possible to generate.

CONCLUSIONS

In conclusion, registrational trials for rare pathogens are difficult, and sometimes impossible, to conduct using traditional statistical criteria. Explicit consideration by the medical, regulatory, and patient community of the ability to correctly approve an effective therapy, the risk of incorrectly approving an ineffective therapy, and the overall number of patients expected to benefit should be used to help define clear success criteria to ensure such studies can be conducted, so that data are available in these important patient groups.

Acknowledgments

Disclaimer. The opinions expressed by the authors do not necessarily reflect the opinion of any of the groups with which they work.

Potential conflicts of interest. A. D. acts as a paid consultant for the Pharmaceutical and Biotechnology Industry. Companies include Achaogen, Allegra, Amicrobe, Amplyx, AN2, Artizan, Cidara, ContraFect, Correvio, Davolterra, Destiny, F2G, Entasis, Geom, GSK, Gyroscope, Humanigen, Kymab, Mironid, Modis, Nabriva, Orca, Pfizer, Phico, Pled, Polyphor, Roche, Scynexis, Sinovent, SNIPR Biome, Spero, TenNor, Transcrip, VenatoRx and Zavante. During the period 2018–2021, J. H. R. reports the following: is/has been Chief Medical Officer & Director, F2G Ltd., Editor-in-Chief, AMR.Solutions, Operating Partner & Consultant, Advent Life Sciences, and Adjunct Professor of Medicine, McGovern Medical School, Houston, TX; has received grant support from Wellcome Trust; sits (or sat) on the scientific advisory boards of Bugworks Research, Inc., Basilea Pharmaceutica, Forge Therapeutics, Inc., Novo Holdings, Roche Pharma Research & Early Development, and Sumitovant; and received consulting fees from Forge Therapeutics, Inc., Innocoll, Vedanta, Progenity, Nosopharm SA, Roivant Sciences, Shionogi Inc., GlaxoSmithKline, and Pfizer Pharmaceuticals. He is currently a

shareholder in AstraZeneca Pharmaceuticals, F2G Ltd., Advent Life Sciences, Zikani Therapeutics, and Bugworks Research, Inc. P. N. is a paid employee of F2G Ltd., has share options in F2G Ltd., and is a shareholder in AstraZeneca Pharmaceuticals. 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. Rex JH. FDA AMDAC – developing single-pathogen agents for *P. aeruginosa* and *A. baumannii*. Available at: <https://amr.solutions/2017/04/13/fda-amdac-developing-single-pathogen-agents-for-p-aeruginosa-and-a-baumannii/>. Accessed 9 June 2022.
2. McKinnell JA, Dwyer JP, Talbot GH, et al. Plazomicin for infections caused by carbapenem-resistant enterobacteriaceae. *N Engl J Med* **2019**; 380:791–3. doi:10.1056/NEJMc1807634.
3. Willan AR, Pinto EM. The value of information and optimal clinical trial design. *Stat Med* **2005**; 24:1791–806. doi:10.1002/sim.2069.
4. Abrahamyan L, Willan AR, Beyene J, et al. Using value-of-information methods when the disease is rare and the treatment is expensive—the example of hemophilia A. *J Gen Intern Med* **2014**; 29(Suppl 3):S767–73. doi:10.1007/s11606-014-2880-3.
5. Pearce M, Hee SW, Madan J, et al. Value of information methods to design a clinical trial in a small population to optimise a health economic utility function. *BMC Med Res Methodol* **2018**; 18:20. doi:10.1186/s12874-018-0475-0.
6. Miller F, Zohar S, Stallard N, et al. Approaches to sample size calculation for clinical trials in rare diseases. *Pharm Stat* **2018**; 17:214–30. doi:10.1002/pst.1848.
7. Rex JH, Talbot GH, Goldberger MJ, et al. Progress in the fight against multidrug-resistant bacteria 2005–2016: modern non-inferiority trial designs enable antibiotic development in advance of epidemic bacterial resistance. *Clin Infect Dis* **2017**; 65:141–6. doi:10.1093/cid/cix246.
8. Rex JH, Eisenstein BI, Alder J, et al. A comprehensive regulatory framework to address the unmet need for new antibacterial treatments. *Lancet Infect Dis* **2013**; 13:269–75. doi:10.1016/S1473-3099(12)70293-1.
9. Chow S-C, Wang H, Shao J. *Sample Size Calculations in Clinical Research*. 2nd ed. New York: Chapman and Hall/CRC, **2007**. doi:10.1201/9781584889830.
10. Entasis Therapeutics Holdings Inc. Form 10-K, page 11. 2020. Available at: <https://www.annualreports.com/Click/28539>. Accessed 9 June 2022.
11. Cidara Therapeutics. Mundipharma and Cidara Therapeutics announce first presentation of results from global Phase 3 restore trial of rezafungin for treatment of candidemia and/or invasive candidiasis demonstrating its positive efficacy and safety profile [press release, 25 April 2022]. Available at: <https://ir.cidara.com/news-releases/news-release-details/mundipharma-and-cidara-therapeutics-announce-first-presentation>. Accessed 9 June 2022.
12. Lagakos SW. Clinical trials and rare diseases. *N Engl J Med* **2003**; 348:2455–6. doi:10.1056/NEJMe030024.
13. Gallin JI, Alling DW, Malech HL, et al. Itraconazole to prevent fungal infections in chronic granulomatous disease. *N Engl J Med* **2003**; 348:2416–22. doi:10.1056/NEJMoa021931.