Pharmaceutical Medicine (2025) 39:73–86
https://doi.org/10.1007/s40290-025-00552-3

CURRENT OPINION

Innovation in the Design of Clinical Trials for Infectious Diseases: Focusing on Patients Over Pathogens

John H. Powers III[1,2] · Robert J. O'Connell[3]

Accepted: 10 February 2025 / Published online: 12 March 2025

Abstract

Much infectious disease research focuses on the interaction of microorganisms and drugs in the laboratory, assuming biological activity of inhibiting organism growth in vitro directly translates to improving patient outcomes in the clinic. Yet in vitro testing does not consider the important role of the human immune system in causing and response to disease. Research shows that patient outcomes are still suboptimal even with disease due to organisms that maintain in vitro susceptibility to currently available drugs. Resources and discussions have focused on "antimicrobial resistance" yet the majority of deaths are with susceptible organisms. Studies of new interventions do not address the questions that patients and clinicians in practice ask in order to improve patient outcomes regardless of causative pathogen in patients who would receive the drugs in the real-world setting. Research in infectious diseases should shift to refocus on improving patient outcomes. This would result in changes in the research questions evaluated, the types of patients enrolled, the comparisons made, the interventions studied, the outcomes evaluated, and the types of statistical evaluations used. In turn this would provide patients and clinicians with better evidence for patient care and justify payment for new interventions.

Plain Language Summary

Infectious disease research has tended to focus on "bugs and drugs" and antimicrobial resistance in Petri dishes in particular. However, focusing on patient outcomes tells a different story, showing that many patients are treated with drugs that kill bacteria in the laboratory but still produce poor patient outcomes, including death. A refocus is needed on patient outcomes over pathogens, with resultant changes in ways in which research is designed, conducted and analyzed in patients with infections. Both patients and doctors need better clinical evidence for new drugs to clearly show that these are beneficial to patients.

1 Introduction

The field of infectious diseases was the first to use research methods such as control groups, blinding/masking, randomization to treatment assignment and inferential statistics [1]. These methodological advances are relatively recent in the long history of medicine, with the first randomized trials held < 80 years ago [2, 3]. These methods are key in allowing causal inferences between interventions and patient outcomes, while minimizing bias and random error, with the ultimate goal of applying the results for patients in clinical practice. The purpose of confirmatory trials is to strengthen the probability of certainty regarding the benefits and harms of new and older interventions. Ultimately, the triad of evidence-based medicine, which includes the best evidence, clinical experience and patient values, enables the generation of evidence that helps patients live longer and live better [4].

Here we will discuss the current state of infectious diseases research and propose ways in which clinical trials may be improved to provide the best evidence for clinicians to aid patients in practice. We will focus on the concepts of types of patients enrolled in studies, the interventions studied, the comparisons evaluated, the outcomes measured and the statistical analyses (PICOS) used in current studies with suggestions ways to improve these moving forward [5].

✉ John H. Powers III
powersjohn@mail.nih.gov

1 Clinical Research Directorate, Frederick National Laboratory for Cancer Research, Frederick, MD, USA

2 George Washington University School of Medicine, Washington, DC, USA

3 Infectious Diseases Clinical Research Program, Bethesda, MD, USA

Key Points

Much infectious disease research focuses on the interaction of bacteria and drugs in the laboratory, assuming this directly translates to improving patient outcomes in clinical care.

Research shows that patient outcomes remain poor, even with disease due to organisms that maintain in vitro susceptibility to current drugs.

Focusing on patient outcomes would result in changes to the research questions evaluated, the types of patients enrolled, comparisons made, interventions studied, outcomes evaluated, and the types of statistical evaluations used.

2 Current State of Research and Evidence from Clinical Trials in Infectious Diseases

2.1 Current Focus on Pathogens Instead of Patients

The current state of research in infectious diseases is characterized by an over-emphasis on pathogenic organisms and small molecule interventions with in vitro biological activity against them—"bugs and drugs" [6] (see Table 1). A large focus and many resources have been placed on the concept of antimicrobial resistance (AMR), with surveillance based on specific organisms and resistance patterns rather than patient outcomes. Various terms have been used to define this idea of focusing on bugs and drugs, including pathogen-focused, pathogen-specific, organism-specific, or pathogen-based development [7, 8]. Rather, research and resulting clinical care should be more patient-focused and agnostic to infecting organisms as to whether such infections are susceptible or resistant to drugs, as long as it improves patient outcomes, for example, host directed therapies which modulate the immune response to infection. By "agnostic" we mean therapies that may benefit patients regardless of the specific infecting organism.

The prevailing notion is that in vitro biological activity defined by minimal inhibitory concentrations (MICs) of drugs against organisms and results from animal models directly translate to understanding benefits and harms of interventions in humans with disease [7, 9]. However, preclinical data that evaluate biological activity are hypothesis generating, not confirmatory evidence of patient benefit. While the germ theory was an important advance in understanding the causal pathway of infections, disease is a result not solely of the presence of organisms, but the host immune response to those organisms and/or their byproducts (e.g., toxins). Indeed, the first therapies for serious infections like pneumonia were host-directed therapies like serum therapy [10]. Past and recent evidence shows that drugs with promising preclinical data routinely fail to demonstrate benefits that exceed harms on patient outcomes [11–13]. This "medicine by MIC" view ignores the contribution to patient outcomes of the human immune system, patient factors and different disease states and severity of disease and misses the larger picture that would come with a broader focus on patients and improving patient outcomes.

2.2 Patients: Enrolling Patients Who Already Have Effective Options

The focus on pathogenic organisms has had several consequences for research and the evidence base for therapeutics in infectious diseases. "Pathogen focused" drug development defines patients by the infecting organism associated with their disease, rather than by whether those patients have an existing need for improved efficacy or less toxic interventions to improve their outcomes.

The current development paradigm for new drugs is based on indirect assumptions rather than direct evidence in patients. Currently, new drugs are studied in trials which indirectly evaluate whether the new drug is more effective than placebo by a comparison with an older effective drug, but do not rule out worse efficacy than current therapies, i.e., "non-inferiority" (NI) [14]. These studies enroll participants who are largely immunocompetent [15] with a common disease for whom current drugs are effective (e.g., urinary tract infections, skin infections, pneumonia or intra-abdominal infections). Yahav et al showed that 90% of trials of new antimicrobials excluded immunocompromised or critically ill patients [13]. However, new drugs are often assumed to be more effective in immunocompromised patients with disease resistant to current therapies [7, 8], although such patients have been excluded from clinical trials for these drugs [15]. Yet this assumption is based on ignoring differences [16] in patient factors, which have a large influence on patient outcomes, noted as early as the 1950s [17]. This also misplaces the meaning of the term "susceptible", as practicing physicians believe that when organisms are categorized as "susceptible" it means there is already evidence that the drug is effective for that type of patient with a specific disease. Instead, the practice of setting "breakpoints" for drug and organism susceptibility has placed an emphasis on in vitro biological activity without requiring evidence of treatment effects in specific types of patients and diseases. For example "epidemiological cut-offs" (ECOFFs) used in Europe are often based on distributions of biological activity against a sample of so-called "wild-type" organisms in vitro rather

Table 1 Current methodologies for studying infectious disease therapies and their challenges in providing evidence to improve patient outcomes and proposed solutions to develop better evidence

Current method for studying infectious disease therapies	Challenges to providing evidence of improving patient outcomes	Solutions to develop better evidence
Focus on micro-organisms rather than improving patient outcomes	Does not necessarily lead to improvements in patient morbidity and/or mortality and focuses on biological effects on organisms in vitro	Refocus research questions on patients and improve patient outcomes
Surveillance based on specific organisms with in vitro resistant profiles	Misses evaluation of impacts of disease from susceptible organisms	Surveillance on patients and outcomes in both antibiotic-susceptible and -resistant disease focusing on baseline patient and laboratory risk factors for poor patient outcomes. Development of definitions of resistance that are clinically relevant and susceptibility criteria moving beyond "breakpoints" that would give clinicians a probability of successful treatment accounting for patient and disease characteristics
Patients enrolled are those who already have effective therapies	Does not focus on patients with greatest need for improved outcomes	Patient enrollment in trials based on probability of lack of effective interventions regardless of causative pathogen to increase feasibility and relevance of clinical trials
Empirical diagnosis	Patients often received inappropriate therapy, either an ineffective drug or a drug not needed at all since no infection is present	Development of rapid point of care diagnostics with demonstrated evidence of improving patient outcomes to better focus administration and proper use of new and older agents
Interventions mostly small molecule drugs with direct inhibition of organism growth	Does not address contribution of host immune system to poor patient outcomes	Interventions beyond small molecule antimicrobials including host directed therapies (including vaccines), microbiome, phage, non-pharmacological interventions (NPIs), etc.
Comparisons usually non-inferiority hypotheses	Need is for interventions with improved efficacy, which is not a research question addressed by non-inferiority hypotheses and raises ethical issues about benefits and harms for those enrolled when no hypothesized non-efficacy benefits (e.g. decreased adverse effects)	Use of randomized trials with superiority hypotheses (including cluster randomization when appropriate for NPIs and vaccines) to improve outcomes for patients when current standards of care are not offering acceptable outcomes, increased use of platform studies and telemedicine when appropriate
Outcomes are usually an indirect measure of patient health such as Clinician Reported Outcomes and biomarkers like negative culture results	Unclear validity of relationship between indirect outcomes such as endpoints and direct patient health status of survival, patient symptoms and patient function in their daily lives	Outcomes based on direct measures of patient's health including survival, patient symptoms and function in their daily lives measured by valid patient reported outcomes (PROs)
Recent use of descriptive statistics in randomized clinical trials evaluating interventions	No clear hypotheses evaluated and an inability to make valid inferences to clinical practice, increased chances of biased/false results	Use of appropriate valid inferential statistics to analyze prespecified hypotheses and appropriate design and evaluation of confirmatory secondary endpoints and subgroup analyses

NPIs non-pharmacological interventions, *PROs* patient-reported outcomes

than evaluating associations with patient outcomes [18]. "Provisional breakpoints" are often determined in the USA using pharmacodynamic (PD) principles and devised prior to the conduct of randomized trials, which appears to be a case of setting the cart before the horse—a drug cannot be determined to be effective in human disease before it has been adequately studied. Recent suggested changes in breakpoints to define susceptibility have been suggested independent of evaluating treatment effects in a valid manner or considering the impact on patient outcomes [19, 20]. In addition the impact on patient outcomes of some types of in vitro resistance remains questionable [21]. Since patients with more resistant organisms are older and sicker, worse outcomes may reflect these patient factors rather than drug efficacy. For instance, a study of *Staphylococcus aureus* blood stream infections showed higher mortality with methicillin-resistant compared to methicillin-susceptible organisms. However, there was no difference in mortality once the investigators controlled for patient factors [22–24].

Pharmacodynamic information is helpful in forming hypotheses regarding which doses to evaluate in randomized trials, but to date there is an absence of evidence that PD dosing predicts benefits for patients in serious diseases like hospitalized pneumonia, despite its widespread use in the drug development process [25]. Claims of the benefits of PD principles are often based on post hoc subgroup analyses from randomized trials; however, these subgroup analyses may be affected by bias as patient factors that occur after randomization may affect drug distribution and clearance (e.g., larger volumes of distribution in the hyperdynamic state of sepsis [26]).

The consequence of the focus on bugs and drugs results in enrolling patients in trials for whom current therapies are on average already effective. This is not the population in whom an unmet need for greater efficacy exists. Focusing on patients in whom existing therapies are effective does not address the primary research question of whether new therapeutics improve outcomes in patients in whom current therapies lack effectiveness, including the population with AMR but also the population with susceptible pathogens in whom current therapies still fail. Ironically despite the pathogen-focus, few people are enrolled in trials who are infected with resistant pathogens [13, 27]. More importantly, recent research shows that 17 of 18 people (94%) with death associated with blood stream infection die with susceptible organisms [28]. The decade old model predicting "more deaths than cancer by 2050" with AMR have not come to pass given decreases or stable numbers to date in AMR and deaths [29]. Recent international data as well as from the US Veterans Administration and the US Department of Defense show low prevalence and decreasing incidence of disease from resistant organisms, and stable or decreasing mortality [30–33]. Since resistance occurs in patients who are sicker, have more co-morbidities and longer duration of hospitalization, attributable mortality is approximately 25% meaning 3 of 4 persons die with, not from, resistant infections [34]. This means that a focus solely on AMR misses the larger burden of patients with disease susceptible to current drug. New drugs with improved in vitro activity against resistant organisms are unlikely to improve outcomes in patients who died with susceptible infections. If AMR is conceived as a "crisis" then deaths in patients with susceptible infections are worthy of clinical and research focus. A return to the pre-antibiotic era is unlikely given the data on decreasing resistance, but in addition volume resuscitation, vasopressor therapy, mechanical ventilation and other forms of ICU care did not exist in the pre-antibiotic era. As seen with Ebola virus disease, better initial resuscitation improved patient outcomes. It is important to take into account that therapies other than antimicrobials, including non-pharmacological therapies, are important in achieving optimal patient outcomes [35].

2.3 Interventions: Sole Focus on "Antibiotics"

The pathogen focus has also resulted in centering on development of "antimicrobial drugs"—small molecule drugs whose mechanism of action is inhibition of organism growth. A review of new antimicrobial drugs showed that most were from existing classes including combinations with beta-lactamase inhibitors. The focus on antimicrobial drugs does not address the role of the disordered immune response in deaths with resistant organisms. The use of systemic antimicrobials also can have adverse effects on the host microbiome. Such effects are not benign and may results in a greater incidence of allergic diseases and decreased efficacy in cancer immunotherapies [36, 37].

2.4 Comparisons: The Problems of Focusing on Randomized Trials with Non-Inferiority Hypotheses and Observational Studies

For clinical trials to provide useful evidence for patients and clinicians, they should address the questions that need answers in clinical practice. The question in the setting of AMR, as well as for patients with susceptible infections in whom current therapies fail, is whether new interventions have improved efficacy in patients who lack effective options. Since current trials enroll patients who already have effective therapies, the comparison in these studies have used the oddly named "non-inferiority" hypotheses [14]. Trials with NI hypotheses are widely misunderstood, misrepresented, misused, and poorly reported [38, 39].

Non-inferiority hypotheses are useful when the benefits of new interventions are not improved efficacy directly but

improved adverse effects or improved convenience (which may eventually improve efficacy for instance from improved adherence) [38]. Non-inferiority hypotheses do not address the question of whether new interventions might improve efficacy over available therapies. Trials with NI hypotheses for efficacy should inherently include evaluating superiority on these other non-efficacy benefits [9]. A review of current infectious disease studies showed a lack of specification of any non-efficacy benefits for new interventions [40, 41].

Despite the implication of the name, NI hypotheses do not evaluate whether a new intervention is "not inferior" to current standard of care. Nor do they evaluate whether a new intervention is "as effective as" or "equivalent" to an older effective intervention [14]. It is statistically impossible to evaluate whether two interventions are identical (zero difference) as this would require an infinite sample size of enrolled participants. Non-inferiority hypotheses address a one-sided question allowing some *loss of efficacy* with a new intervention—not "not inferior"—compared to an older intervention already shown to be effective under the conditions of the planned trial in trade off for non-efficacy benefits.

The need for some benefits to balance potential harms for trial participants is a requirement in the Belmont Report on the ethical conduct of research involving humans [42]. Therefore, it is unethical to enroll participants in NI trials with no proposed non-efficacy benefits assuming hypothetical benefits in unstudied types of future patients in whom the current control drug might lose effectiveness due to AMR [43].

This extrapolation of demonstration of NI in patients who have effective therapies to superiority in unstudied patients who lack effective therapies is scientifically invalid as well as unethical. The results of NI trials only apply to patients in whom the control drug is effective and meet the enrollment criteria for NI studies. Studies show the types of patients with AMR are older, sicker, have more co-morbidities and more prior therapies than patients infected with susceptible organisms [44]. These patients are routinely excluded from NI trials [15]. The results of trials with cefiderocol showed NI in patients with complicated urinary tract infections who lacked resistance but the same drug increased mortality in a randomized trial of patients with resistant organisms [45]. Patient outcomes can vary widely not just by infecting organism but by patient characteristics even when infected with the same organism. The Belmont report requires that the types of patients who bear the risks of research should be those who also obtain its benefits [42].

The inability to use placebo to avoid loss of efficacy of available therapies in a serious and life-threatening disease is not a justification to use NI hypotheses. The Declaration of Helsinki points out that patients should not be randomized to placebo *or any drug less effective than the best available one* when effective therapy exists and when patients might be exposed to irreparable harm [46]. Diseases classified as "serious and life-threatening" by regulatory agencies are categorized in this way specifically because of their potential to cause irreparable harm. Since patients might often choose to experience more adverse effects or greater inconvenience in exchange for greater efficacy in life-threatening diseases, the trade-offs in NI hypotheses do not make sense in acute life-threatening bacterial infections [47]. There may be no second chance at rescue therapy if the initial agents fail.

Furthermore, the amount of potential loss of efficacy chosen before trial initiation even in settings where participants would not be exposed to irreparable harm is supposed to be clinically inconsequential to patients [48, 49]. However, studies show the amount of loss of efficacy allowed in trials with NI hypotheses has increased over time from 10 to 20% (1 in 5 participants potentially have worse outcomes) [12]. A recent survey of potential research participants found that over half would not agree to participate in a NI trial of a new infectious disease therapeutic that allowed a 10% loss of efficacy regardless of other benefits [50]. The goal of NI trials is not to assure that the new drug is simply better than placebo but to maintain a large proportion of the treatment effect of the control drug [49], especially in serious diseases, balanced by demonstration of non-efficacy benefits. A new drug as much as 20% *more* effective than current therapies would be considered a major advance; therefore, it is challenging to understand how a potential loss of efficacy as large as 20% could be considered clinically inconsequential.

The notion that the point estimate of the treatment difference in an NI trial is "the most likely true value" in a *single trial*, while the upper and lower bounds of the confidence interval (CI) are "unlikely", is a common misunderstanding of frequentist CIs, confusing them with Bayesian credibility intervals [51, 52]. Frequentist CIs are based on a theoretical replication of a study over multiple repetitions, with 95% of the studies on average containing the true difference somewhere within the bounds of the CI, and 5% of studies giving erroneous results by not containing the true mean with the CI bounds of the study. The point estimates will be closer to the true mean on average across many studies, but in a single trial the CIs may not even include the true mean difference between the interventions [53].

Furthermore, research on informed consent forms for infectious disease trials with NI hypotheses showed that participants are not informed that the primary purpose of the study allows the new intervention to be potentially less effective [41]. The notion that NI trial results might on occasion demonstrate superior results for the new intervention is not relevant at the time of consent as participants can only know trial hypotheses, not the results, prior to trial initiation [54]. Using patients who have effective therapies as means to

an end for regulatory approval for marketing while offering them no potential benefits in the NI study is unethical [55].

The results of studies with NI hypotheses often are misinterpreted [14, 56, 57]. Results with 95% CIs that include no difference do not rule out some amount of harm. Smaller trials with fewer participants give less precision on estimates of effects but do not mean a new drug "still might be a little better". For example, a new drug studied in a NI trial with results with a point estimate of the difference between the new drug minus the older drug on the primary outcomes gives −1.0%, with a 95% CI of −4.0 to 3.0%. This result meets the criteria for demonstration of NI using a −10% loss of efficacy chosen prior to the trial. However, these results do not mean "the new drug still might be as much as 3% better". Since NI is a one-sided question, the results only address how much *less* effective the new intervention might be in comparison to the older effective agent as shown by the lower bound of the CI. This does not rule out potential harm of up to 4% less efficacy compared to the older effective control drug. A trial, which enrolls fewer participants, that gives the same point estimate result of −1.0% might have 95% CIs of −9.0 to 8.0%. Again, this meets the criteria for NI using a 10% loss of efficacy but also demonstrates that smaller trials give less precise results and wider CIs. This does not mean that the new drug becomes up to 8% more effective merely by studying fewer patients. The larger the loss of efficacy chosen for the new drug prior to the trial, the smaller the number of participants is needed to meet the chosen criteria for NI. For example, the US Food and Drug Administration (FDA) allowed the single study that formed the basis for approval for sulbactam-durlobactam to be up to 20% less effective than the control drug, colistin, in order to allow enrollment of fewer patients in the study [58]. Therefore, there is an incorrect incentive in NI trials to allow larger losses of efficacy prior to the trial with the goal of enrolling fewer participants (at the same time decreasing the database for adverse effects), while erroneously interpreting the accompanying lack of precision as "the new drug might be better" due to wider CIs. International guidance points out the loss of effect of the control drug should be clinically inconsequential to patients, and the loss of effect should not be chosen solely based on decisions to decrease the number of enrolled participants in the trial [49].

Lastly, trials using NI hypotheses are more prone to bias even if they are randomized and blinded [38]. These are more likely to bias towards no difference, giving a false positive conclusion of NI. One study showed over 90% of trials with NI hypotheses were "successful" but it is unclear how use of the NI hypothesis should result in a higher rate of positive findings compared to superiority trials [59]. Unfortunately, the use of NI hypotheses has become the default trial design in infectious diseases, regardless of the relevant clinical question. Funders have urged investigators to use NI even when they attempt to design trials using superiority hypotheses [60]. The conclusions of a successful NI trial do not answer which intervention is better so should be used preferentially—the answer clinicians are seeking—but rather only demonstrates two interventions are 'interchangeable" within the bounds set ahead of the trial. Broad conclusions of "showing non-inferiority" are meaningless without understanding the amount of efficacy loss allowed prior to the trial, the results compared to that choice, and the endpoints evaluated.

Trials in HIV provide an example of using NI hypotheses appropriately. A trial evaluating less frequent dosing (every 8 weeks compared with every 4 weeks) with long-acting injectable agents was designed to address the non-efficacy benefit of greater convenience—which in a chronic disease may promote greater adherence, which in turn may be improve efficacy [61]. These trials do not expose patients to irreversible harm as those who experience virological failure were rescued with oral agents and only allowed a hypothesis of a small amount of loss of efficacy of 4% on virological failure at 152 weeks. The results of this study are also show that the results were statistically inferior for the 8-week regimen on the primary endpoint, and simultaneously non-inferior since the amount of efficacy loss was below the amount chosen by investigators prior to the study (point estimate of 1.7% less effective on virological outcomes for 8 weeks vs 4 weeks with 95% CI of 0.1 to 3.3% less effective while excluding zero). Of interest, patients were not queried for their views on this amount of loss of efficacy, yet this shows that non-inferior does not mean "equal".

There has also been an increased emphasis on the use of so-called "real world evidence" framed as comparisons from non-randomized, often retrospective, observational studies, for example, post-marketing studies of new antibiotics in patients with AMR and test-negative study designs in evaluating vaccine efficacy [62, 63]. The advent of randomization in the 1940s was genius—using chance to counter chance and addressing the major issue of unmeasured confounding, which no amount of adjustment in observational studies can address. In observational studies differences remain in who does and does not receive an intervention [64]. Often, data used for billing are not accurate representations of patients' health status. Methods used to adjust for various biases often assume "no unmeasured confounding", an assumption that is unverifiable at best and erroneous at worst [65]. The larger number of patients in an observational study, the larger the effect of bias on study results. Therefore, having larger observational databases—even if the data are accurate—does not solve the challenges of bias and confounding—it worsens them. A recent study showed that giving the same observational dataset testing the same hypothesis to 73 different investigative groups resulted in a wide variety of answers from that same dataset, some

in opposite directions of effect, showing the challenges in the stubbornness of unmeasured biases not only in the data themselves, but in the choices investigators make in analysis as well as other unexplained sources of variance [66]. These choices are often opaque to readers of medical studies. Even when accurately reported, the impact of those choices on biasing study results is often unclear. For example, despite randomized trial evidence showing worse mortality with cefiderocol including a numerical increase of 34% mortality in patients with *Acinetobacter* infections, an observational study claimed a survival benefit in these same types of patients [67]. Yet only 11% of patients in the observational study received monotherapy and only certain types of patients were included in the analysis, making this type of evidence challenging for clinical decision making in the face of randomized trial evidence to the contrary. Observational studies can address questions like who is at risk of disease but their use in evaluating treatment effects of medical interventions is not an advance, but a regression to methods that often gave erroneous results in the past.

2.5 Outcomes: What Happens to the Bug or What Happens to Patients?

Since the goal of prescribing or taking a medical intervention is to help patients to live longer and to live better, clinical trials—especially those in acute diseases—should directly measure survival, patient symptoms, and patient function in their daily lives. However, most clinical trials in infectious diseases use indirect "surrogate" endpoints, usually comprising clinician-reported outcomes (such as clinician decisions to prescribe another non-study drug to treat the disease) or biomarkers such as negative cultures, body temperature, antibody concentrations and physical signs of disease [12]. The term "clinical cure" is often poorly defined with the specific factors measured, how they are combined, how they are evaluated and the criteria for success left to the discretion of individual investigators. Endpoints cannot be well defined nor reliable when evaluated in this way. A recent review showed that 21 FDA guidance in 27 different infectious disease indications recommended use of surrogate endpoints, either alone or combined with other outcomes [68]. A recent review of the evidence for new antimicrobials showed that none used direct patient-reported outcomes of symptoms or patient function as endpoints [12]. Fever—defined as body temperature measurement—is a commonly used and particularly poor surrogate endpoint. Fever is not on the causal pathway of disease but rather the disease causes fever. This is important because variables not on the causal pathway can be affected by interventions that do not impact the disease, such as antipyretics affecting body temperature but not the disease outcomes. While there is some logic to using surrogate endpoints in chronic diseases like HIV where direct patient outcomes may take years to occur, there is little reason to use surrogate endpoints in acute diseases like skin infections, pneumonia, abdominal infections and urinary tract infections, which are the most commonly studied diseases.

2.6 Statistics: Describing a Sample or Making Inferences?

With the use of biostatistics in medical studies, it is important to address random error and evaluate the play of chance on study results. The advent of inferential statistical methods was crucial to improving the causal assessments of interventions and outcome. Linked to the use of inferential statistics is clearly stating a hypothesis that the study is evaluating. Descriptive statistics are used to outline the results in a single study sample with no attempt to infer what those outcomes mean for larger populations, such as describing the numbers of patients who experience pneumonia in a single hospital. In contrast, inferential statistics analyze a sample from a population attempting to make inferences regarding that larger population from which the sample was drawn. Recently, infectious disease trials have been designed with no hypothesis and "descriptive statistics only" to "gain experience" with new therapeutics [45]. The challenge with no hypothesis is one can interpret the results in whatever manner one chooses after the study is completed after looking at the data. Ironically, results in these trials are presented using *p*-values and CIs, which are themselves methods of inferential statistics [52, 53]. The stated goal for using descriptive statistics is to speed up the marketing of new drugs, yet these methods are not advances but a reversion to methods abandoned almost a century ago. Inappropriate statistical methods do not meet the goal of providing better evidence for patients and clinicians. Furthermore, it is unclear how trials with no hypotheses and insufficient statistical methods meet the ethical goal of being scientifically valid and socially valuable, which is a requirement for all research involving human participants [55]. A claimed unmet need is a reason to conduct a study, yet its design and conduct should be sufficient to be able to determine whether it provides evidence that the need can be met. Scientific and ethical review boards should give more scrutiny to the use of appropriate design methods in infectious disease trials.

2.7 Summarizing the Current State of Research and Evidence

The result of the pathogen focus has been a lack of evidence of benefit with new therapeutics, mostly antibiotics, with increasing costs while ignoring the larger group of patients with poor outcomes infected with susceptible pathogens [13]. The claimed benefits of new therapeutics have been assumed based on the notion that therapeutics are similarly effective across different types of patients as long as they are infected with the same "susceptible" organism in vitro, without direct evidence in those patients in whom the drugs would be used in practice. This view ignores the impact of patient factors on disease outcomes. Three recent reviews evaluate the evidence for new therapeutic agents in infectious diseases in the USA, encompassing a total of 23 new interventions (excluding vaccines) [11–13, 27]. While the need for new drugs is based on claims related to AMR, few patients enrolled in the studies had disease due to resistant organisms.

3 Moving Forward—Suggestions to Improve the Evidence Base

3.1 Refocusing on Patients Instead of Pathogens

The unmet need in infectious diseases is to improve patient outcomes regardless of infecting pathogen (see Table 1). The sole reason AMR is important is because it is associated with worse patient outcomes. While there has been much discussion on AMR, a focus on patients and patient outcomes reveals a different story. First, resistance is often defined by how many drugs or drug classes might lack effectiveness in vitro (multi-drug resistance [MDR]), a question that is not relevant to practicing clinicians or patients if effective drugs remain. The concept of "breakpoints"—dichotomizing organisms into susceptible versus intermediate or resistant—is less helpful to clinicians than determining continuous probabilities of treatment success based on in vitro testing. Evaluating probabilities of treatment success should be based on valid assessments of treatment effects in specific types of patients with specific diseases, as case reports of drug failures or in vitro data alone are not able to assess outcomes attributable to treatments.

A more clinically relevant definition of "resistance" is how many effective drugs might remain (difficult to treat resistance), which is better linked to patient outcomes than MDR [69, 70]. Assessing resistance in this way shows that the greater burden in terms of number of deaths is associated with patients infected with organisms susceptible to currently available drugs. If AMR is considered a public health crisis, then deaths with susceptible pathogens deserve at least as much if not more consideration. Refocusing on patients instead of pathogens has several downstream consequences for both research and clinical care.

3.2 Patients

Understanding the greater burden of disease from susceptible organisms means we should focus surveillance on patients with poor outcomes in specific disease syndromes, regardless of infecting pathogen. A focus on "priority pathogens" based on in vitro susceptibility will miss this greater burden of disease. More research is needed to be able to identify patients most at risk of treatment failure with currently available drugs even if the infecting organism is categorized as "susceptible". This would allow enrollment criteria for trials of new interventions to focus on this group of patients as well. The number of patients who might benefit even if there were evidence of superior efficacy with new drugs targeted only to resistant pathogens is small and as noted above does not seem to be increasing in developed countries. However, the overall number with poor outcomes remains large [71].

3.3 Interventions

The types of new interventions evaluated should expand beyond small molecule drugs with direct acting antimicrobial activity. First, this should include rapid point of care diagnostics to enable better application of current drugs. Recent research shows that lack of appropriate empirical therapy has a substantial impact on patient outcomes [72]. Infectious diseases are unique in their reliance on empiricism instead of directly applied diagnostics. Evaluation of new diagnostics should demonstrate their impact on patient outcomes rather than solely focusing on detecting more organisms. For example, use of rapid point of care diagnostics in tuberculosis and sexually transmitted infections has been shown to improve patient outcomes in randomized trials [73, 74].

Second, the types of interventions evaluated should include prevention as well as treatment either with active and/or passive immunization, which could be agnostic to specific organisms. Non-pharmacological interventions (NPIs) should also be evaluated to determine their ability to prevent initiation and spread of disease. History has shown an improvement in mortality with infections even before the first antibiotic was developed, in large part due to infection control, and clean water and food [75]. New interventions should include evaluation of other forms of direct-acting antimicrobial therapies such as phage, which

is currently being targeted to patients with "resistant" infections, yet may also benefit patients with "susceptible" infections [76], either as sole therapy or in addition to current small molecule drug antimicrobials. Another reason for poor outcomes in patients with susceptible organisms is a disordered immune response, and immunotherapies should also be evaluated. In the COVID-19 pandemic one of the first effective interventions was host immune response modification with corticosteroids [77]. Studies have also demonstrated improved survival with corticosteroids in severe pneumonia [78]. Microbiome therapy has also hypothesized benefits on patient outcomes through host immune interactions [79].

3.4 Comparisons

The greatest need for patients in whom current therapies fail is interventions with improved efficacy. The valid assessment of such interventions should be in randomized trials with superiority hypotheses [80]. Superiority trials can be designed as either a direct comparison of a new intervention to current standard of care or as add-on trials to current standard of care compared to standard of care alone, as commonly done in oncology. The claim that superiority trials are "unethical" is again based on bug-centered focus that one should not randomize patients to a drug to which an organism is resistant. However, if patients in the control group are receiving current best standard of care, the resistance pattern of an organism does not determine the ethics of a trial. The use of NI hypotheses should be reserved for interventions whose hypothesized benefits are on non-efficacy benefits like decreased adverse effects or improved convenience in settings where patients would not be exposed to increased harm due to loss of efficacy relative to the effective older intervention. The NI hypotheses do not evaluate improved efficacy of new interventions.

The COVID-19 pandemic and Ebola virus disease outbreaks saw an increase in usage of platform trials, which allow comparison of multiple interventions against a common control group [77, 81]. Rather than evaluating the benefits and harms of a single intervention, platform trials address the question of which of several interventions is best for patients compared to current standard of care. Platform trials require more planning and would require greater cooperation between drug sponsors, funders and regulators [77, 82, 83]. Telemedicine has been utilized more often in efforts to expand geographic diversity of trial participants and ease trial participation [84].

Randomized trials and real-world evidence are not incompatible [62]. Improving the choices in enrollment criteria and outcome measures in pragmatic trials can reflect real-world practice that would improve the ability to use clinical trial evidence in real-world settings [85, 86].

3.5 Outcomes

Outcome measures in observational studies used to define endpoints in randomized trials should focus on outcomes that are most important to patients, including survival, symptoms and function in their daily lives [87]. Survival is not "confounded" in randomized trials, as the process of randomization gives a similar distribution of the probability of death from other causes besides the studied interventions between the test and control group [80]. Since patient factors highly influence outcomes in serious diseases, this reinforces the importance of randomization, as unmeasured confounding still remains in many observational studies [64, 88]. Furthermore, death can occur from drug adverse effects as well as lack of efficacy. Finally, other outcomes cannot be measures in patients who do not survive. The goal of randomized trials is to evaluate outcomes attributable to the interventions studied, not attributable to the disease.

Endpoints in trials could be improved by developing and choosing them based on (1) what is measured, (2) how the measure is performed and (3) how the endpoint data are analyzed. What is measured in trials should be clearly defined and include all-cause mortality as well as valid measures of patient function and symptoms. There is a lack of evidence that the surrogate endpoint of acute clearance of organisms from cultures reflects patient outcomes [89–91]. The measurements should be performed using valid patient-reported outcome (PRO) instruments as done with COVID-19 and other viral respiratory diseases [92–95]. There is a need to develop valid PROs for other diseases including skin infections, pneumonia and urinary tract infections [96]. Patient-reported outcomes can measure various outcomes including disease symptoms, adverse effects as well as more distal health-related quality of life. In other settings where patients cannot self-report, valid clinician-reported outcomes may be used [97]. The desirability of outcome ranking (DOOR) is an analysis method that ranks outcomes from the most important to the least important [98]. However, DOOR alone does not define an endpoint and it is only as strong as its measured components and the validity of the measurements made. The impact on survival related to infection goes beyond the acute period due to the continued impact of inflammation, therefore the timing of endpoints requires longer follow-up to evaluate this impact [99, 100].

3.6 Statistics

The analysis of analytical research studies comparing interventions should utilize appropriate inferential statistics [101, 102]. The post-World War II advent of statistical methods occurred because of the errors and biases inherent in observational research before that time. Descriptive statistics are appropriate to describe the results in a sample without

extrapolating or comparing them to another group, e.g., how many patients in a hospital develop pneumonia in one year? In descriptive studies, there are often no stated hypotheses. Inferential statistics are appropriate when "inferring" the results from a studied sample to the larger population from which the sample was taken. This includes randomized trials when testing hypotheses in a sample of patients with a given disease are studied and the results are inferred to reflect the benefits and harms in the larger population with that disease with similar characteristics to enrolled patients. Descriptive statistics are not appropriate when attempting to evaluate the effect of interventions in the population from which the sample was selected. Trials analyzed in this way having no pre-stated hypotheses make conclusions after examining the data, which can bias the results. Furthermore, it is unclear how the use of descriptive statistics either aids patients or aids drug development, as unclear or biased results may increase patient harm. The use of Bayesian statistics, using prior evidence of drug effectiveness in the analysis of current trial results, is possible but often the prior evidence upon which to base such analyses is often absent for new therapies in infectious diseases or changes in medical care make prior evidence less applicable to current care. The utility of adaptive methods depends upon what is adapted and how, and requires advance planning [102]. For example, adding and eliminating study interventions in so-called "platform trials", which evaluate multiple interventions against a single control group, was used during the COVID-19 pandemic and in Ebola trials but required much advanced planning [77, 103]. Secondary endpoint and subgroup analyses in trials should be carefully planned and more than "pre-specification" is needed to make valid inferences [104–106]. Testing multiple subgroups can increase the probability of false positive results without planning and adjusting statistical methods in the protocol to account for multiple comparisons. Subgroup analyses are particularly suspect when they give conclusions that differ from the primary trial results [107]. The use of methods like propensity scoring to compare similar patients who do and do not receive interventions in observational studies is increasing, yet this method is valid only when unmeasured confounding is unlikely, which is rare. Propensity scoring attempts to adjust for baseline variables but does not account for post-baseline differences in care and other interventions received [65]. Observational data can be useful in evaluating which patients are most at risk of disease and which outcomes occur in which types of patients, but randomized trials are still needed to evaluate unbiased effects of interventions.

Analyses in randomized superiority trials should be conducted according to the groups to which patients were randomized, i.e., the intention to treat (ITT) population analysis. The ITT analysis maintains the protection that randomization provides from selection bias, ensuring the patients analyzed are similar in baseline prognostic factors other than the interventions studied. However, a challenge with studies with NI hypotheses comes from the fact that there is no optimal analysis population. The ITT analyses may bias towards no difference especially if there are large amounts of missing data, but per-protocol analyses often exclude patients based on factors after randomization and may also give biased results [108].

4 Conclusions

The history of infectious diseases research coincided with the progress of experimentation to obtain valid results on the benefits and harms of medical interventions for clinical care [1]. In recent years, infectious disease research has been based primarily on issues regarding pathogenic organisms instead of patient outcomes. This has resulted in enrolling patients in trials who already have effective interventions, a singular focus on trials with non-inferiority hypotheses evaluating small molecule antimicrobial drugs, using outcomes that are not direct measures of patient benefit analyzed with descriptive statistics. This has increased the quantity but not the quality of available interventions. To advance the field of infectious disease research and help patients, the focus should be on patients and patient outcomes and trials should include patients in whom current therapies are not effective, evaluating superior hypotheses in randomized trials on direct patient outcomes using appropriate inferential statistical methods. Refocusing on patients instead of pathogens will improve patient outcomes and the value of clinical care [109, 110].

Declarations

Funding Open access funding provided by the National Institutes of Health. This project has been funded in whole or in part with federal funds from the National Institutes of Health Clinical Center and the National Cancer Institute, National Institutes of Health, under Contract No. 75N91019D00024 Task Order No. 75N91019F00130. 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.

Conflict of Interest The authors have no conflicts of interest for this manuscript.

Ethics Approval Not applicable.

Consent to Participate Not applicable.

Consent for Publication Not applicable.

Availability of Data and Materials Not applicable.

Code Availability Not applicable.

Authors' Contributions JP initial planning, conducting literature search, and drafting, editing. RJO contributing to conceptualization, editing manuscript.

References

1. Kaptchuk TJ, Kerr CE. Commentary: unbiased divination, unbiased evidence, and the patulin clinical trial. Int J Epidemiol. 2004;33(2):247–51.
2. Medical Research Council. Streptomycin treatment of pulmonary tuberculosis. Br Med J. 1948;2(4582):769–82.
3. Chalmers I, Clarke M. Commentary: the 1944 patulin trial: the first properly controlled multicentre trial conducted under the aegis of the British Medical Research Council. Int J Epidemiol. 2004;33(2):253–60.
4. Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. BMJ. 1996;312(7023):71–2.
5. Saaiq M, Ashraf B. Modifying "Pico" Question into "Picos" model for more robust and reproducible presentation of the methodology employed in a scientific study. World J Plast Surg. 2017;6(3):390–2.
6. Boucher HW, Talbot GH, Bradley JS, Edwards JE, Gilbert D, Rice LB, Scheld M, Spellberg B, Bartlett J. Bad bugs, no drugs: no ESKAPE! An update from the infectious diseases Society of America. Clin Infect Dis. 2009;48(1):1–12.
7. US Food and Drug Administration. Antibacterial therapies for patients with unmet medical need for the treatment of serious bacterial diseases: Guidance for Industry. https://www.fda.gov/files/drugs/published/Antibacterial-Therapies-for-Patients-With-an-Unmet-Medical-Need-for-the-Treatment-of-Serious-Bacterial-Diseases.pdf.
8. US Food and Drug Administration. Antibacterial therapies for patients with an unmet medical need for the treatment of serious bacterial diseases. 2017.
9. Doshi P. Did the FDA break its own rules in approving the antibiotic Recarbrio? BMJ. 2023;381: p1048.
10. Finland M. Serum therapy in the treatment of lobar pneumonia. N Engl J Med. 1930;202(26):1244–7.
11. Deak D, Outterson K, Powers JH, Kesselheim AS: Progress in the fight against multidrug-resistant bacteria? A review of U.S. food and drug administration-approved antibiotics, 2010–2015. Ann Intern Med 2016.
12. Mitra-Majumdar M, Powers JH, Brown BL, Kesselheim AS. Evidence at time of regulatory approval and cost of new antibiotics in 2016–19: cohort study of FDA approved drugs. BMJ Med. 2022;1(1): e000227.
13. Yahav D, Tau N, Shepshelovich D. Assessment of data supporting the efficacy of new antibiotics for treating infections caused by multidrug-resistant bacteria. Clin Infect Dis. 2021;72(11):1968–74.
14. Powers JH. Noninferiority and equivalence trials: deciphering "similarity" of medical interventions. Stat Med. 2008;27(3):343–52.
15. Kuzucan A, Powers JH, Doshi P. Antibiotics approved for marketing in populations specifically excluded from premarketing trials, 1999–2018. Mayo Clin Proc. 2020;95(12):2699–703.
16. Dowling HF. Twixt the cup and the lip. J Am Med Assoc. 1957;165(6):657–61.
17. Fleming TR, Powers JH. Issues in noninferiority trials: the evidence in community-acquired pneumonia. Clin Infect Dis. 2008;47(Suppl 3):S108-120.
18. Kahlmeter G, Turnidge J. 'How To: ECOFFs—the why, the how and the don'ts of EUCAST epidemiological cutoff values'—Author's response. Clin Microbiol Infect. 2022;28(7):1030–1.
19. Tamma PD, Powers JH. Do patient data really support the clinical and laboratory standards institute recommendation for lowering third-generation cephalosporin interpretive breakpoints? Clin Infect Dis. 2013;57(4):624–5.
20. Tamma PD, Wu H, Gerber JS, Hsu AJ, Tekle T, Carroll KC, Cosgrove SE. Outcomes of children with enterobacteriaceae bacteremia with reduced susceptibility to ceftriaxone: do the revised breakpoints translate to improved patient outcomes? Pediatr Infect Dis J. 2013;32(9):965–9.
21. Nuermberger E, Bishai WR. The clinical significance of macrolide-resistant Streptococcus pneumoniae: it's all relative. Clin Infect Dis. 2004;38(1):99–103.
22. Holmes NE, Turnidge JD, Munckhof WJ, Robinson JO, Korman TM, O'Sullivan MV, Anderson TL, Roberts SA, Warren SJ, Gao W, et al. Vancomycin AUC/MIC ratio and 30-day mortality in patients with *Staphylococcus aureus* Bacteremia. Antimicrob Agents Chemother. 2013;57(4):1654–63.
23. Holmes NE, Turnidge JD, Munckhof WJ, Robinson JO, Korman TM, O'Sullivan MV, Anderson TL, Roberts SA, Warren SJ, Gao W, et al. Vancomycin minimum inhibitory concentration, host comorbidities and mortality in Staphylococcus aureus bacteraemia. Clin Microbiol Infect. 2013;19(12):1163–8.
24. Holmes NE, Turnidge JD, Munckhof WJ, Robinson JO, Korman TM, O'Sullivan MV, Anderson TL, Roberts SA, Gao W, Christiansen KJ, et al. Antibiotic choice may not explain poorer outcomes in patients with Staphylococcus aureus bacteremia and high vancomycin minimum inhibitory concentrations. J Infect Dis. 2011;204(3):340–7.
25. Lux LJ et al. Pharmacokinetic/Pharmacodynamic measures for guiding antibiotic treatment for hospital-acquired pneumonia. http://www.ncbi.nlm.nih.gov/pubmed/25577937.
26. Abdul-Aziz MH, Alffenaar JC, Bassetti M, Bracht H, Dimopoulos G, Marriott D, Neely MN, Paiva JA, Pea F, Sjovall F, et al. Antimicrobial therapeutic drug monitoring in critically ill adult patients: a position paper. Intensive Care Med. 2020;46(6):1127–53.
27. Yahav D, Shepshelovich D, Tau N. Cost analysis of new antibiotics to treat multidrug-resistant bacterial infections: mind the gap. Infect Dis Ther. 2021;10(1):621–30.
28. Lawandi A, Kadri S, Powers J. Focusing on antimicrobial resistant infections—Are we missing the forest for the trees and the patients for pathogens? Front Antibiotics. 2023. https://doi.org/10.3389/frabi.2023.1329081.
29. Diallo OO, Baron SA, Dubourg G, Chaudet H, Halfon P, Camiade S, Comte B, Joubert S, Francois A, Seyral P, et al. Major discrepancy between factual antibiotic resistance and consumption in South of France: analysis of 539,037 bacterial strains. Sci Rep. 2020;10(1):18262.

30. Baquero F. Threats of antibiotic resistance: an obliged reappraisal. Int Microbiol. 2021;24(4):499–506.
31. Abat C, Fournier PE, Jimeno MT, Rolain JM, Raoult D. Extremely and pandrug-resistant bacteria extra-deaths: myth or reality? Eur J Clin Microbiol Infect Dis. 2018;37(9):1687–97.
32. Abat C, Rolain JM, Dubourg G, Fournier PE, Chaudet H, Raoult D. Evaluating the clinical burden and mortality attributable to antibiotic resistance: the disparity of empirical data and simple model estimations. Clin Infect Dis. 2017;65(suppl_1):S58–63.
33. Liao JX, Appaneal HJ, Menon A, Lopes V, LaPlante KL, Caffrey AR. Decreasing antibiotic resistance trends nationally in gram-negative bacteria across United States veterans affairs medical centers, 2011–2020. Infect Dis Ther. 2023;12(7):1835–48.
34. Antimicrobial Resistance Collaborators. Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. Lancet. 2022;399(10325):629–55.
35. Fischer WA, Wohl DA. Combining vaccines, optimised supportive care, and therapeutics for Ebola virus disease increases survival. Lancet Infect Dis. 2024;24(6):560–2.
36. von Itzstein MS, Gonugunta AS, Sheffield T, Homsi J, Dowell JE, Koh AY, Raj P, Fattah F, Wang Y, Basava VS et al. Association between antibiotic exposure and systemic immune parameters in cancer patients receiving checkpoint inhibitor therapy. Cancers (Basel). 2022;14(5):1327.
37. Lovegrove MC, Geller AI, Fleming-Dutra KE, Shehab N, Sapiano MRP, Budnitz DS. US emergency department visits for adverse drug events from antibiotics in children, 2011–2015. J Pediatric Infect Dis Soc. 2019;8(5):384–91.
38. Mauri L, D'Agostino RB. Challenges in the design and interpretation of noninferiority trials. N Engl J Med. 2017;377(14):1357–67.
39. Piaggio G, Elbourne DR, Pocock SJ, Evans SJ, Altman DG, Group C. Reporting of noninferiority and equivalence randomized trials: extension of the CONSORT 2010 statement. JAMA. 2012;308(24):2594–604.
40. Doshi P, Spence O, Kuzucan A, Powers JH. Communication of nonefficacy benefits of new drugs approved on the basis of noninferiority trials alone: cohort study of FDA and sponsor communication, 2011–2017. JAMA Intern Med. 2019;179(5):719–21.
41. Doshi P, Hur P, Jones M, Albarmawi H, Jefferson T, Morgan DJ, Spears PA, Powers JH. Informed consent to study purpose in randomized clinical trials of antibiotics, 1991 through 2011. JAMA Intern Med. 2017;177(10):1–8.
42. Ethical Principles and Guidelines for the Protection of Human
43. Ethical Principles and Guidelines for the Protection of Human Subjects of Research: the Belmont Report. http://www.hhs.gov/ohrp/humansubjects/guidance/belmont.html.
44. Garattini S, Bertele' V. Non-inferiority trials are unethical because they disregard patients' interests. Lancet. 2007;370(9602):1875–7.
45. Safdar N, Maki DG. The commonality of risk factors for nosocomial colonization and infection with antimicrobial-resistant Staphylococcus aureus, enterococcus, gram-negative bacilli, Clostridium difficile, and Candida. Ann Intern Med. 2002;136(11):834–44.
46. Powers JH. Scientific evidence, regulatory decision making, and incentives for therapeutics in infectious diseases: the example of cefiderocol. Clin Infect Dis. 2021;72(12):e1112–4.
47. Declaration of Helsinki. http://www.wma.net/en/30publications/10policies/b3/.
48. Doshi P, Spence O, Powers JH. Noninferiority trials. N Engl J Med. 2018;378(3):304.
49. International Conference on Harmonization. Statistical principles for clinical trials E9. http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E9/Step4/E9_Guideline.pdf.
50. International Conference on Harmonization. Choice of control group and related issues in clinical trials. http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E10/Step4/E10_Guideline.pdf.
51. Morlock R, Rose J, Powers JH. Perspectives of US adults on antimicrobial trials with noninferiority designs. JAMA Netw Open. 2023;6(5): e2316297.
52. Belia S, Fidler F, Williams J, Cumming G. Researchers misunderstand confidence intervals and standard error bars. Psychol Methods. 2005;10(4):389–96.
53. Hoekstra R, Morey RD, Rouder JN, Wagenmakers EJ. Robust misinterpretation of confidence intervals. Psychon Bull Rev. 2014;21(5):1157–64.
54. Cumming G, Finch S. Inference by eye: confidence intervals and how to read pictures of data. Am Psychol. 2005;60(2):170–80.
55. Beecher HK. Ethics and clinical research. N Engl J Med. 1966;274(24):1354–60.
56. Freedman B. Scientific value and validity as ethical requirements for research: a proposed explication. IRB. 1987;9(6):7–10.
57. Odem-Davis K, Fleming TR. A simulation study evaluating biocreep risk in serial non-inferiority clinical trials for preservation of effect. Stat Biopharm Res. 2015;7(1):12–24.
58. Bai AD, Komorowski AS, Lo CKL, Tandon P, Li XX, Mokashi V, Cvetkovic A, Kay VR, Findlater A, Liang L, et al. Methodological and reporting quality of noninferiority randomized controlled trials comparing antibiotic therapies: a systematic review. Clin Infect Dis. 2021;73(7):e1696–705.
59. US Food and Drug Administration: Sulbactam-durlobactam (Xacduro) Intergrated Review; 2024. https://www.accessdata.fda.gov/drugsatfda_docs/nda/2023/216974Orig1s000TOC.cfm.
60. Li Y, He Y, Sheng Y, Wang K, Wang J, Huang J, Yang J, Zheng Q. Systematic evaluation of non-inferiority and equivalence randomized trials of anti-infective drugs. Expert Rev Anti Infect Ther. 2013;11(12):1377–89.
61. Drekonja DM, Trautner B, Johnson JR. Effect of 7 vs 14 days of antibiotics among afebrile men with urinary tract infection-reply. JAMA. 2021;326(20):2080–1.
62. Overton ET, Richmond G, Rizzardini G, Thalme A, Girard PM, Wong A, Porteiro N, Swindells S, Reynes J, Noe S, et al. Long-acting cabotegravir and rilpivirine dosed every 2 months in adults with human immunodeficiency virus 1 type 1 infection: 152-week results from ATLAS-2M, a randomized, open-label, phase 3b, noninferiority study. Clin Infect Dis. 2023;76(9):1646–54.
63. Sherman RE, Anderson SA, Dal Pan GJ, Gray GW, Gross T, Hunter NL, LaVange L, Marinac-Dabic D, Marks PW, Robb MA, et al. Real-world evidence—What is it and what can it tell us? N Engl J Med. 2016;375(23):2293–7.
64. Shrank WH, Patrick AR, Brookhart MA. Healthy user and related biases in observational studies of preventive interventions: a primer for physicians. J Gen Intern Med. 2011;26(5):546–50.
65. Bosco JL, Silliman RA, Thwin SS, Geiger AM, Buist DS, Prout MN, Yood MU, Haque R, Wei F, Lash TL. A most stubborn bias: no adjustment method fully resolves confounding by indication in observational studies. J Clin Epidemiol. 2010;63(1):64–74.
66. Freemantle N, Marston L, Walters K, Wood J, Reynolds MR, Petersen I. Making inferences on treatment effects from real world data: propensity scores, confounding by indication, and other perils for the unwary in observational research. BMJ. 2013;347: f6409.
67. Breznau N, Rinke EM, Wuttke A, Nguyen HHV, Adem M, Adriaans J, Alvarez-Benjumea A, Andersen HK, Auer D, Azevedo F, et al. Observing many researchers using the same data and hypothesis reveals a hidden universe of uncertainty. Proc Natl Acad Sci USA. 2022;119(44): e2203150119.

68. Epling BP, Powers JH 3rd. Cefiderocol and the need for higher-quality evidence: methods matter for patients. Antimicrob Agents Chemother. 2022;66(8): e0076622.
69. Hey SP, Kesselheim AS, Patel P, Mehrotra P, Powers JH. US food and drug administration recommendations on the use of surrogate measures as end points in new anti-infective drug approvals. JAMA Intern Med. 2019. https://doi.org/10.1001/jamainternmed.2019.5451.
70. Kadri SS, Adjemian J, Lai YL, Spaulding AB, Ricotta E, Prevots DR, Palmore TN, Rhee C, Klompas M, Dekker JP, et al. Difficult-to-treat resistance in gram-negative bacteremia at 173 US hospitals: retrospective cohort analysis of prevalence, predictors, and outcome of resistance to all first-line agents. Clin Infect Dis. 2018. https://doi.org/10.1093/cid/ciy378.
71. Kadri SS, Lai YLE, Ricotta EE, Strich JR, Babiker A, Rhee C, Klompas M, Dekker JP, Powers JH, Danner RL, et al. External validation of difficult-to-treat resistance prevalence and mortality risk in gram-negative bloodstream infection using electronic health record data from 140 US hospitals. Open Forum Infect Dis. 2019;6(4): ofz110.
72. Strich JR, Warner S, Lai YL, Demirkale CY, Powers JH 3rd, Danner RL, Kadri SS. Needs assessment for novel Gram-negative antibiotics in US hospitals: a retrospective cohort study. Lancet Infect Dis. 2020;20(10):1172–81.
73. Kadri SS, Lai YL, Warner S, Strich JR, Babiker A, Ricotta EE, Demirkale CY, Dekker JP, Palmore TN, Rhee C, et al. Inappropriate empirical antibiotic therapy for bloodstream infections based on discordant in-vitro susceptibilities: a retrospective cohort analysis of prevalence, predictors, and mortality risk in US hospitals. Lancet Infect Dis. 2021;21(2):241–51.
74. Peter JG, Zijenah LS, Chanda D, Clowes P, Lesosky M, Gina P, Mehta N, Calligaro G, Lombard CJ, Kadzirange G, et al. Effect on mortality of point-of-care, urine-based lipoarabinomannan testing to guide tuberculosis treatment initiation in HIV-positive hospital inpatients: a pragmatic, parallel-group, multicountry, open-label, randomised controlled trial. Lancet. 2016;387(10024):1187–97.
75. Gaydos CA, Ako MC, Lewis M, Hsieh YH, Rothman RE, Dugas AF. Use of a rapid diagnostic for chlamydia trachomatis and neisseria gonorrhoeae for women in the emergency department can improve clinical management: report of a randomized clinical trial. Ann Emerg Med. 2019;74(1):36–44.
76. Hemminki E, Paakkulainen A. The effect of antibiotics on mortality from infectious diseases in Sweden and Finland. Am J Public Health. 1976;66(12):1180–4.
77. Tu Q, Pu M, Li Y, Wang Y, Li M, Song L, An X, Fan H, Tong Y. Acinetobacter baumannii phages: past, present and future. Viruses. 2023;15(3):673.
78. Normand ST. The RECOVERY platform. N Engl J Med. 2021;384(8):757–8.
79. Dequin PF, Meziani F, Quenot JP, Kamel T, Ricard JD, Badie J, Reignier J, Heming N, Plantefève G, Souweine B, et al. Hydrocortisone in severe community-acquired pneumonia. N Engl J Med. 2023;388(21):1931–41.
80. Zheng D, Liwinski T, Elinav E. Interaction between microbiota and immunity in health and disease. Cell Res. 2020;30(6):492–506.
81. Altman DG, Bland JM. Statistics notes. Treatment allocation in controlled trials: why randomise? BMJ. 1999;318(7192):1209.
82. Mulangu S, Dodd LE, Davey RT, Tshiani Mbaya O, Proschan M, Mukadi D, Lusakibanza Manzo M, Nzolo D, Tshomba Oloma A, Ibanda A, et al. A randomized, controlled trial of Ebola virus disease therapeutics. N Engl J Med. 2019;381(24):2293–303.
83. Simon R. Critical review of umbrella, basket, and platform designs for oncology clinical trials. Clin Pharmacol Ther. 2017;102(6):934–41.
84. Dodd LE, Freidlin B, Korn EL. Platform trials - beware the noncomparable control group. N Engl J Med. 2021;384(16):1572–3.
85. Tarolli CG, Andrzejewski K, Zimmerman GA, Bull M, Goldenthal S, Auinger P, O'Brien M, Dorsey ER, Biglan K, Simuni T. Feasibility, reliability, and value of remote video-based trial visits in Parkinson's disease. J Parkinsons Dis. 2020;10(4):1779–86.
86. Califf RM. Pragmatic clinical trials: emerging challenges and new roles for statisticians. Clin Trials. 2016;13(5):471–7.
87. Ford I, Norrie J. Pragmatic trials. N Engl J Med. 2016;375(5):454–63.
88. Granholm A, Anthon CT, Kjær MN, Maagaard M, Kaas-Hansen BS, Sivapalan P, Schjørring OL, Andersen LW, Mathiesen O, Strøm T, et al. Patient-important outcomes other than mortality in contemporary ICU trials: a scoping review. Crit Care Med. 2022;50(10):e759–71.
89. Hartz A, He T, Wallace R, Powers J. Comparing hormone therapy effects in two RCTs and two large observational studies that used similar methods for comprehensive data collection and outcome assessment. BMJ Open 2013, 3(7).
90. Montezuma-Rusca JM, Powers JH, Follmann D, Wang J, Sullivan B, Williamson PR. Early fungicidal activity as a candidate surrogate endpoint for all-cause mortality in cryptococcal meningitis: a systematic review of the evidence. PLoS One. 2016;11(8): e0159727.
91. Fleming TR, Powers JH. Biomarkers and surrogate endpoints in clinical trials. Stat Med. 2012;31(25):2973–84.
92. Powers JH. Microbiologic surrogate end points in clinical trials of infectious diseases: example of acute otitis media trials. Pharmacotherapy. 2005;25(12 Pt 2):109S-123S.
93. Curran D, Matthews S, Cabrera ES, Pérez SN, Breva LP, Rämet M, Helman L, Park DW, Schwarz TF, Melendez IMG, et al. The respiratory syncytial virus prefusion F protein vaccine attenuates the severity of respiratory syncytial virus-associated disease in breakthrough infections in adults ≥60 years of age. Influenza Other Respir Viruses. 2024;18(2): e13236.
94. Satram S, Ghafoori P, Reyes CM, Keeley TJH, Birch HJ, Brintziki D, Aldinger M, Alexander E, Lopuski A, Sarkis EH, et al. Assessment of symptoms in COMET-ICE, a phase 2/3 study of sotrovimab for early treatment of non-hospitalized patients with COVID-19. J Patient Rep Outcomes. 2023;7(1):92.
95. Richard SA, Epsi NJ, Lindholm DA, Malloy AMW, Maves RC, Berjohn CM, Lalani T, Smith AG, Mody RM, Ganesan A, et al. COVID-19 patient-reported symptoms using FLU-PRO plus in a cohort study: associations with infecting genotype, vaccine history, and return to health. Open Forum Infect Dis. 2022;9(7): ofac275.
96. Yu J, Powers JH, Vallo D, Falloon J. Evaluation of efficacy endpoints for a phase IIb study of a respiratory syncytial virus vaccine in older adults using patient-reported outcomes with laboratory confirmation. Value Health. 2020;23(2):227–35.
97. Powers JH, Howard K, Saretsky T, Clifford S, Hoffmann S, Llorens L, Talbot G. Patient-reported outcome assessments as endpoints in studies in infectious diseases. Clin Infect Dis. 2016;63(Suppl 2):S52-56.
98. Powers JH, Patrick DL, Walton MK, Marquis P, Cano S, Hobart J, Isaac M, Vamvakas S, Slagle A, Molsen E, et al. Clinician-reported outcome assessments of treatment benefit: report of the ISPOR clinical outcome assessment emerging good practices task force. Value Health. 2017;20(1):2–14.
99. Evans SR, Rubin D, Follmann D, Pennello G, Huskins WC, Powers JH, Schoenfeld D, Chuang-Stein C, Cosgrove SE, Fowler VG, et al. Desirability of Outcome Ranking (DOOR) and response adjusted for days of antibiotic risk (RADAR). Clin Infect Dis. 2015. https://doi.org/10.1093/cid/civ495.
100. Le T, Kinh NV, Cuc NTK, Tung NLN, Lam NT, Thuy PTT, Cuong DD, Phuc PTH, Vinh VH, Hanh DTH, et al. A trial of

itraconazole or amphotericin B for HIV-associated talaromycosis. N Engl J Med. 2017;376(24):2329–40.
101. Shankar-Hari M, Rubenfeld GD. Understanding long-term outcomes following sepsis: implications and challenges. Curr Infect Dis Rep. 2016;18(11):37.
102. Guetterman TC. Basics of statistics for primary care research. Fam Med Community Health. 2019;7(2): e000067.
103. Emerson SS, Fleming TR. Adaptive methods: telling "the rest of the story." J Biopharm Stat. 2010;20(6):1150–65.
104. Dodd LE, Proschan MA, Neuhaus J, Koopmeiners JS, Neaton J, Beigel JD, Barrett K, Lane HC, Davey RT. Design of a randomized controlled trial for ebola virus disease medical countermeasures: PREVAIL II, the Ebola MCM study. J Infect Dis. 2016;213(12):1906–13.
105. Freemantle N. Interpreting the results of secondary end points and subgroup analyses in clinical trials: should we lock the crazy aunt in the attic? BMJ. 2001;322(7292):989–91.
106. Assmann SF, Pocock SJ, Enos LE, Kasten LE. Subgroup analysis and other (mis)uses of baseline data in clinical trials. Lancet. 2000;355(9209):1064–9.
107. O'Neill RT. Secondary endpoints cannot be validly analyzed if the primary endpoint does not demonstrate clear statistical significance. Control Clin Trials. 1997;18(6):550–6 (**discussion 561-557**).
108. Powers JH, Ross DB, Lin D, Soreth J. Linezolid and vancomycin for methicillin-resistant Staphylococcus aureus nosocomial pneumonia: the subtleties of subgroup analyses. Chest. 2004;126(1):314–5 (author reply 315-316).
109. Evans S, Rubin DB, Powers JH, Follmann D. Analysis populations in anti-infective clinical trials: whom to analyze? Stat Commun Infect Dis. 2018. https://doi.org/10.1515/scid-2017-0002.
110. Porter ME. What is value in health care? N Engl J Med. 2010;363(26):2477–81.