

OPINION

Open Access

Optimising trial designs to identify appropriate antibiotic treatment durations



Koen B. Pouwels^{1,2,3*} , Mo Yin^{4,5}, Christopher C. Butler^{6,7}, Ben S. Cooper^{4,8}, Sarah Wordsworth^{1,6,9}, A. Sarah Walker^{6,9,10†} and Julie V. Robotham^{2,6†}

Abstract

Background: For many infectious conditions, the optimal antibiotic course length remains unclear. The estimation of course length must consider the important trade-off between maximising short- and long-term efficacy and minimising antibiotic resistance and toxicity.

Main body: Evidence on optimal treatment durations should come from randomised controlled trials. However, most antibiotic randomised controlled trials compare two arbitrarily chosen durations. We argue that alternative trial designs, which allow allocation of patients to multiple different treatment durations, are needed to better identify optimal antibiotic durations. There are important considerations when deciding which design is most useful in identifying optimal treatment durations, including the ability to model the duration–response relationship (or duration–response ‘curve’), the risk of allocation concealment bias, statistical efficiency, the possibility to rapidly drop arms that are clearly inferior, and the possibility of modelling the trade-off between multiple competing outcomes.

Conclusion: Multi-arm designs modelling duration–response curves with the possibility to drop inferior arms during the trial could provide more information about the optimal duration of antibiotic therapies than traditional head-to-head comparisons of limited numbers of durations, while minimising the probability of assigning trial participants to an ineffective treatment regimen.

Keywords: Antimicrobial resistance, Design, Randomised trial, Duration of therapy, Antibiotics, Bayesian, Frequentist

Background

Bacteria are increasingly able to resist antibiotic treatment, resulting in increased morbidity, deaths and costs worldwide [1–5]. Antibiotic use is an important driver of the development and spread of antimicrobial resistance [6]. Selective pressure can be reduced by minimising antibiotic prescribing for conditions for which antibiotics are often unnecessary [7–10]. Moreover, the duration of antibiotic courses can often be reduced without significantly compromising cure rates [11–15]. Shortening antibiotic duration can have a large impact on reducing exposure of bacteria to antibiotics, including bacteria

carried asymptotically [14]. Antibiotics are by far the most prescribed drugs for children, with more than 60 million systemic antibiotics dispensed annually in the US outpatient setting alone [16], and are amongst the most frequently prescribed drugs for adults [17].

An important challenge is that, for many infectious conditions, the optimal antibiotic course length remains unclear. Courses should be long enough to treat infections effectively, yet short enough to reduce the incidence of side effects and the development and spread of antibiotic resistance. The continued need for antibiotics can be assessed with daily reviews for inpatients [18]; however, in practice, such reviews are not always performed or acted upon and antibiotics are often continued in order to complete currently recommended course durations [19]. Further, in the outpatient or primary care setting, continued assessment of patients initiated on antibiotics is not feasible [14]. Therefore, it is especially

* Correspondence: koen.pouwels@npdh.ox.ac.uk

†A. Sarah Walker and Julie V. Robotham contributed equally to this work.

¹Health Economics Research Centre, Nuffield Department of Population Health, University of Oxford, Oxford, UK

²Modelling and Economics Unit, National Infection Service, Public Health England, London, UK

Full list of author information is available at the end of the article



important to have strong evidence about optimal treatment durations in these settings.

Because observational studies comparing different antibiotic durations are potentially confounded by unmeasured patient factors influencing the need for prolonged treatment, evidence about the optimal treatment duration should, where possible, come from randomised controlled trials (RCTs). However, antibiotic durations for several infections managed in primary care, such as prostatitis, are not guided by RCT evidence on optimal treatment duration [20, 21]. Where treatment durations have been compared in RCTs, in most cases, two treatment durations were selected for comparison, both of which lacked a clear scientific rationale [11, 12]. Whilst RCTs designed in this way can be useful, there are disadvantages to this approach.

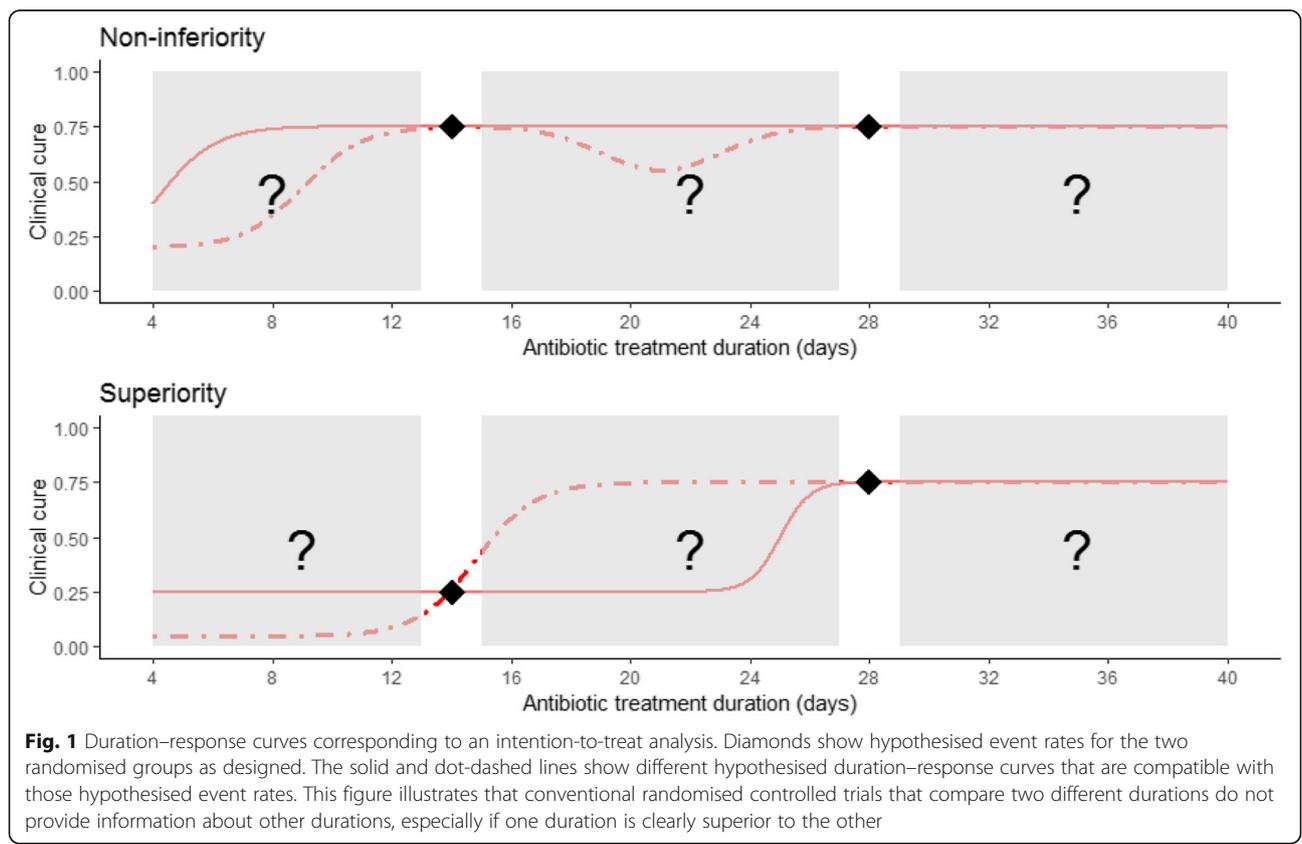
Herein, we discuss the main issue with conventional two-arm trial designs, how to assess the ‘optimal’ antibiotic treatment duration, four alternative trial designs that can estimate much needed duration–response relationships (subsequently denoted duration–response curves), and which of these designs has the most desirable properties.

Issue with conventional designs

Historically, RCTs have an experimental and comparator arm, or two contrasting experimental arms [22]. An

issue with conventional two-arm trials is that they are unlikely to identify optimal treatment durations, potentially leading to suboptimal clinical practice. An approach that is more likely to identify optimal treatment durations is the modelling of duration–response curves.

In the specific example of prostatitis, we could design a conventional RCT comparing, for example, treatment durations of 14 versus 28 days. Depending on whether the trial is designed to show superiority or non-inferiority, the trial answers the question ‘is 14 days of antibiotic treatment for prostatitis as good as/worse/better than 28 days of treatment?’ (Fig. 1). However, this does not answer the more important question of ‘what is the optimal antibiotic treatment duration for prostatitis?’ The dot-dashed line in the top panel could occur if there is some non-compliance with the shorter duration as randomised, because patients are not cured (e.g. may still have persisting minor symptoms, which might then relapse without further antibiotic treatment) at the end of their assigned duration (and hence more patients end up receiving the standard duration (28 days) despite being randomised to a shorter duration). In practice, not all patients will require the same duration and, at the population level, the proportion of patients that are not cured will likely decrease with increasing assigned duration, as will the proportion with non-compliance, creating the dot-dashed line in the top panel.



How to estimate the ‘optimal duration’?

Selecting the optimum treatment duration depends on the outcomes that are deemed important, which are often measures of cure (treatment effectiveness), either in the short-term or medium- to long-term such as prevention of relapse/recurrence. Secondary outcomes usually relate to side effects and, sometimes, to the development of resistance. The fact that cure rates can generally be hypothesised to increase with duration until reaching an asymptote creates a delicate balance between maximising efficacy and minimising adverse consequences.

The Desirability of Outcome Ranking/Response Adjusted for Duration of Antibiotic Risk (DOOR/RADAR) trial design has recently been proposed as a method to formally combine clinical outcomes and treatment duration into a single composite outcome [23]. However, in its ranking, this approach implicitly assumes that the shorter of two durations is beneficial when other patient or clinical outcomes are identical [24]. This unverified strong assumption could lead to demonstration of non-inferiority using DOOR/RADAR when conventional trial designs may show that shorter durations are not non-inferior [24].

In situations where the optimal decision regarding the best treatment depends on various endpoints, maximising a utility function (or minimising a loss function), a decision-theoretic Bayes (or full Bayesian) approach provides an intuitive solution [25, 26]. A recent Bayesian response-adaptive randomised trial evaluating the usefulness of gepotidacin for the treatment of patients with Gram-positive acute bacterial skin and skin structure infections used a utility function to determine the optimal treatment dose [25]. The dose–response-for-cure rate was modelled using a normal dynamic linear model with the parameter evolution described by a Gaussian random walk, while the dose–response-for-the-discontinuation rate was modelled with a two-parameter logistic model assuming a monotonic change [25]. The cure rate component and the treatment discontinuation component were combined multiplicatively to yield the final utility [25]. An advantage of using a utility function is that the trade-offs between the different components are made explicit and quantified. This approach provides the answer we really want to know, namely ‘what is the optimal treatment duration taking into account the trade-offs between efficacy and safety and antibiotic resistance development?’ However, given the difficulty in devising a generally acceptable utility function and the computational complexities, the decision-theoretic Bayes approach using a utility function is rarely used [26]. A problem with applying the decision-theoretic approach in medicine is that there are many decision-makers at different stages, including policy-makers, physicians and patients, who likely have different opinions and utility

functions [27]. For example, different individuals may not agree that the cure and the discontinuation rates can be combined multiplicatively, and therefore also question the validity of the response-adaptive changes in the trial.

Given these and other difficulties with implementing a decision-theoretic approach [28], it may be more practical to model separate duration–response curves for efficacy, antibiotic resistance and toxicity during the trial, and combine the information from the different duration–response curves with additional information, such as costs or estimated longer term influence on resistance, into a decision analytic framework [29]. Optimal durations can then be assessed for various prior opinions and utility functions of the different stakeholders.

The optimal treatment duration may differ depending on host- or pathogen-specific characteristics. Formally, this can be tested by including different subgroup-specific interaction terms in models relating duration to response [29]. This could allow stratified medicine, enabling different optimal durations to be identified depending on key patient characteristics.

Main characteristics of alternative RCT designs

Recently, various groups have suggested that fixed or adaptive trials could be used to assess which treatment duration should be recommended [29, 30]. Herein, we discuss multi-arm designs to model the duration–response curve, including (1) a Bayesian response-adaptive randomisation (RAR) design [31]; (2) a play-the-winner design [30, 32]; (3) a fixed duration design [29]; and (4) a drop-the-loser design.

All four designs could be used in combination with flexible regression modelling strategies to model the duration–response curve, such as fractional polynomials, or in the case of frequent reassessment of the duration–response curve, normal dynamic linear models [25, 29, 33]. To account for uncertainty about the structural form of the duration–response curve generating the data, model averaging can be used [34, 35].

The main differences between the four alternative designs and the typical two-arm frequentist randomised trial are listed in Table 1.

The Bayesian RAR design allows allocation of a greater proportion of future patients to treatment durations that have performed well at the interim analysis based on posterior predictive probabilities [31]. Unless the posterior predictive probability is too low (arm should be dropped) or sufficiently high (the arm may be selected for the next phase of testing or selected as the optimal treatment), the updated randomisation probability is typically proportional to the predictive probability of success for the experimental relative to the control arm [28].

The play-the-winner design includes an option to continue with the shortest duration that has a posterior

Table 1 Main characteristics of conventional two-arm and alternative multi-arm duration trial designs

	Two-arm duration design	Bayesian multi-arm, response-adaptive randomisation design	Multi-arm, play-the-winner design	Multi-arm, fixed duration design	Multi-arm, drop-the-loser duration design
Obtain information on treatment durations	Only two considered durations	Can model the entire duration–response curve, including estimates for durations that were not used	Can model the entire duration–response curve, including estimates for durations that were not used	Can model the entire duration–response curve, including estimates for durations that were not used	Can model the entire duration–response curve, including estimates for durations that were not used
Randomisation probabilities	Fixed	Variable and possibility to drop arms	Fixed, but with possibility to select shortest duration with high probability of being non-inferior for comparison with standard duration	Fixed	Fixed, but with possibility to drop arms
Higher chance for patients to be randomised to better-performing treatment arm	No	Yes	Yes	No	No, but poor performing arms can be dropped

predictive probability (or a frequentist test statistic) above a pre-defined threshold compared to the standard duration, based on the assumption that shorter durations will be better in terms of risk of antibiotic resistance and toxicity.

The fixed duration design has been recently proposed by Quartagno et al. [29]. In contrast to the other designs, this is not an adaptive trial design, but focusses on gaining sufficient information to accurately model the ‘duration–response’ relationship.

The drop-the-loser design uses stopping boundaries to determine whether experimental treatments should be dropped early by comparing each to the standard (assumed maximum duration) treatment. At interim analysis, it can be decided to drop clearly inferior treatment arms compared to the standard duration based on Bayesian posterior predictive probabilities [26], or based on other predefined stopping criteria using a frequentist test-statistic [32, 36].

Which design has the most desirable properties?

Important considerations when designing a randomised trial to identify optimal antibiotic treatment durations include (1) the ability to accurately estimate the duration–response curve, (2) the ability to minimise the risk of bias, (3) the possibility to drop poorly performing arms, and (4) statistical efficiency.

Estimating duration–response curves

A major benefit of estimating a duration–response curve is that the effects of durations not included in the trial can also be estimated provided that there is sufficient data from neighbouring durations. This applies to all four alternative RCT designs considered here.

However, a potential issue with adaptive designs that preferentially assign patients to better performing arms (RAR and play-the-winner designs) is that this may hamper proper evaluation of the complete duration–response

curve due to an insufficient number of patients receiving different durations. One could prevent this issue by assigning patients preferentially to informative treatment durations, i.e. durations that would increase the precision in an area of the curve, or by setting a threshold to the maximum imbalance in randomisation probabilities. Nevertheless, in practice, it may be more feasible to use designs with fixed randomisation probabilities (fixed duration design), potentially with the option to drop arms that are clearly inferior to the standard duration (drop-the-loser design).

Subgroup-specific duration–response curves could be obtained by including interaction terms for pre-specified subgroups such as immunocompromised patients who may require longer antibiotic therapy. With RAR (all designs except the fixed duration design), changes in allocation ratios can theoretically be based on the duration–response curve within subgroups in the presence of a subgroup effect, making the trial statistically less efficient. However, it is difficult to identify subgroup effects during a trial given the lower power to detect them, and these are usually only assessed at the end of a trial. Therefore, designs which drop arms or allocate proportionately fewer patients to some arms on a population level basis (i.e. using results from the trial as a whole), may end up without sufficient information to assess whether the optimal duration varies across important subgroups.

Risk of bias

An important challenge that applies to all alternative RCT designs comparing multiple antibiotic treatment durations is the difficulty in blinding clinicians and patients. Where a perfectly matching placebo is available and instructions are provided about the order of taking preparations, blinding is theoretically possible, yet, in practice, such a placebo is difficult/expensive to make. Therefore, duration RCTs are often open-label. When using an open-label design that preferentially allocates

patients to specific durations with better outcomes (RAR design), clinicians will be able to determine, during the trial, that these durations are associated with better outcomes, thereby increasing the risk of allocation concealment (selection) bias. This knowledge can change which patients get randomised in the trial and how endpoints will subsequently be assessed. The other designs all reduce the risk of selection bias because clinicians cannot alter the selection of patients based on observed changes in allocation probabilities for these designs.

It is often cautioned that calendar time trends – which are common with infectious diseases – may introduce bias when using RAR [30, 37]. However, one can take advantage of the fact that randomisation probabilities are not constantly changing with most RAR designs. A calendar time-stratified analysis, with equal randomisation probabilities within each stratum, eliminates potential time-trend bias [38]. A larger sample size is needed with such stratified analyses, but it is important to avoid trying to gain small improvements in efficiency at the cost of introducing bias [38]. While the fixed duration design is not vulnerable to time trends due to its design, the RAR and play-the-winner design require a less efficient calendar time-stratified analysis to avoid this type of bias. When considering a drop-the-loser design one should avoid comparison of patients assigned to the dropped duration with patients that were randomised to other arms after dropping the clearly inferior arm to avoid this bias. However, this may not be problematic given that there was already enough information to deem the duration clearly inferior.

An issue encountered with all antibiotic duration trials is potential non-compliance. Non-compliance can provide a distorted picture of the efficacy of treatment durations when performing an intention-to-treat (ITT) analysis (Fig. 1). In an ITT analysis, patients are analysed according to their assigned duration, regardless of whether they actually received that duration. ITT analyses provide unbiased estimates of effectiveness, i.e. the real-world impact of the intention to receive one versus another duration, assuming that the type of non-compliance that occurred in the trial would generalise outside the trial. In situations where non-compliance reduces the difference in treatment received between two arms being compared, an ITT analysis is not conservative for a test of non-inferiority. The effect of non-compliance, which is likely not completely random, can be taken into account using instrumental variable approaches and/or g-methods as described in more detail by Berry et al. [27] and Hernan et al. [39].

Dropping poorly performing arms

For all designs, including fixed trial designs, continuous response monitoring for serious and unexpected adverse

events or lack of efficacy of certain durations by an independent data monitoring committee can ensure that patients are protected from being randomised to an unsafe arm [40]. For adaptive designs, futility stopping criteria are defined at the planning stage. This can be done for both frequentist and Bayesian trials and would provide statistical rules to help the data monitoring committee decide whether an arm should be dropped [27, 28, 32]. After dropping an arm, follow-up will continue for patients assigned to this duration. The advantage of having the option to drop poorly performing arms (drop-the-loser design) is that it potentially reduces the number of patients allocated to unfavourable antibiotic durations. This is not only ethically desirable, but may also convince more patients to participate in a trial.

Statistical efficiency and sample size

In the recent proposal for the fixed duration design [29], simulations showed that a sample size of 500 patients divided into 5–7 equidistant arms was sufficient to estimate the duration–response curve within a 5% error margin in 95% of the simulations, suggesting that a trial using similar methodology is feasible in practice [29]. Similar simulations focusing on the numbers needed to estimate duration–response curves for the other designs do not yet exist. In general, using standard pairwise comparisons, the more arms included, the greater the sample size, but it is not clear that such pairwise comparisons are ideal for determining optimal treatment duration.

The main reason for the increasing interest in adaptive trial designs (all designs, except the fixed duration design) may be that, under some circumstances, adaptive designs are statistically more efficient than fixed trial designs [32, 37, 38, 41, 42]. However, as mentioned earlier, if patients are preferentially allocated to the best performing arms, the precision around other durations of the duration–response curve will be reduced [29]. In addition, as discussed above, to prevent bias due to time trends, stratified analysis is recommended [38], which requires a larger sample size than the potentially biased un-stratified analysis, the latter often being used in simulations comparing response-adaptive and fixed duration designs [38, 41, 42].

The verdict

Given the considerations laid out above, the fixed duration and the drop-the-loser duration designs theoretically have the most potential to identify optimal antibiotic treatment durations. These designs (1) are less vulnerable to allocation concealment bias than the RAR design; (2) are not vulnerable (fixed duration) or are less vulnerable (drop-the-loser) to time-trend bias compared to the RAR or play-the-winner designs; (3) are

not associated with the important logistical challenges often accompanying adaptive trials that allow for changes in the allocation ratios (play-the-winner and RAR designs) [32, 43, 44]; and (4) are more likely than the RAR and play-the-winner designs to have sufficient numbers of patients in each arm and/or subgroup at the end of the trial to estimate the complete duration–response curve with sufficient precision, and hence enable evaluation of the potential for important differences in the optimal duration within specific subgroups.

A potential advantage of the drop-the-loser design over the fixed duration design is that the former can drop duration arms that are clearly inferior versus the standard (maximum) duration based on formal statistical analysis. This may ethically be more acceptable by reducing the number of patients allocated to inferior treatment durations.

Although we have only provided theoretical considerations regarding these four designs, we urge the research community to consider developing, testing and applying alternative trial designs that can identify optimal treatment durations, including sample size calculations.

Extensions

Whilst we have focussed on antibiotic duration, evidence supporting doses of many commonly used antibiotics is similarly scarce, and similar methods could also be used to optimise dose. In practice, particularly in primary care, different durations may well be completely equivalent in terms of acute recovery, yet rare but important complications may vary with different durations. Very large numbers would need to be randomised to estimate ‘duration–response curves’ for these rare outcomes, potentially as co-primary endpoints, or incorporated in a decision analytic framework together with other outcomes [45]. Finally, in the context of changing patterns of resistance or access to care, for example, the optimal duration for any specific indication today may not be optimal tomorrow. A platform duration trial, which allows for the dropping and addition of arms, could be a solution to providing continuously relevant evidence [46], and would also enable different durations of different drugs to be compared.

Conclusion

There is a clear need for more evidence on optimal antibiotic treatment durations. Multi-arm designs that estimate duration–response curves have a much higher probability of finding the optimal duration for different conditions and patient populations than conventional two-arm RCTs. More research into the properties of alternative RCT designs that can estimate

duration–response curves are needed, as well as actual applications of such designs to better identify optimal antibiotic treatment durations. Strengthening the evidence on antibiotic treatment duration is critical in guiding antibiotic stewardship and reducing harm from antibiotic resistance and adverse drug effects.

Abbreviations

DOOR/RADAR: Desirability of Outcome Ranking/Response Adjusted for Duration of Antibiotic Risk; ITT: Intention-to-treat; RAR: Response-adaptive randomisation; RCT: Randomised controlled trial

Acknowledgements

Not applicable.

Authors' contributions

KBP, MY, CCB, BSC, SW, ASW and JVR all contributed to the conception, writing and editing of this opinion article. All authors have approved the final version.

Funding

ASW is supported by the NIHR Oxford Biomedical Research Centre and core support to the MRC Clinical Trials Unit at UCL [MC_UU_12023/22 and MC_UU_12023/29], and is an NIHR Senior Investigator.

Availability of data and materials

Not applicable.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Health Economics Research Centre, Nuffield Department of Population Health, University of Oxford, Oxford, UK. ²Modelling and Economics Unit, National Infection Service, Public Health England, London, UK. ³Department of Health Sciences, Global Health, University Medical Centre Groningen, University of Groningen, 9713, GZ, Groningen, The Netherlands. ⁴Mahidol Oxford Tropical Medicine Research Unit, Faculty of Tropical Medicine, Mahidol University, Bangkok, Thailand. ⁵Division of infectious disease, University Medicine Cluster, National University Hospital, Singapore, Singapore. ⁶The National Institute for Health Research Health Protection Research Unit in Healthcare Associated Infections and Antimicrobial Resistance at the University of Oxford, Oxford, UK. ⁷Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford, UK. ⁸Nuffield Department of Medicine, Centre for Tropical Medicine and Global Health, University of Oxford, Oxford, UK. ⁹National Institute for Health Research Oxford Biomedical Research Centre, John Radcliffe Hospital, University of Oxford, Oxford, UK. ¹⁰MRC Clinical Trials Unit at University College London, London, UK.

Received: 30 January 2019 Accepted: 20 May 2019

Published online: 21 June 2019

References

1. de Kraker MEA, Jarlier V, Monen JCM, Heuer OE, van de Sande N, Grundmann H. The changing epidemiology of bacteraemias in Europe: trends from the European antimicrobial resistance surveillance system. *Clin Microbiol Infect.* 2013;19:860–8.
2. Whittles L, White P, Paul J, Didelot X. Epidemiological trends of antibiotic resistant gonorrhoea in the United Kingdom. *Antibiotics.* 2018;7(3):60.
3. Shrestha P, Cooper BS, Coast J, Oppong R, Do Thi Thuy N, Phodha T, et al. Enumerating the economic cost of antimicrobial resistance per antibiotic consumed to inform the evaluation of interventions affecting their use. *Antimicrob Resist Infect Control.* 2018;7:98.

4. Naylor NR, Atun R, Zhu N, Kulasabanathan K, Silva S, Chatterjee A, et al. Estimating the burden of antimicrobial resistance: a systematic literature review. *Antimicrob Resist Infect Control*. 2018;7:58.
5. Roope LSJ, Smith RD, Pouwels KB, Buchanan J, Abel L, Eibich P, et al. The challenge of antimicrobial resistance: what economics can contribute. *Science*. 2019;364(6435):eaau4679. <https://doi.org/10.1126/science.aau4679>.
6. Chatterjee A, Modarai M, Naylor NR, Boyd SE, Atun R, Barlow J, et al. Quantifying drivers of antibiotic resistance in humans: a systematic review. *Lancet Infect Dis*. 2018;18:e368–78.
7. Pouwels KB, Dolk FCK, Smith DRM, Robotham JV, Smieszek T. Actual versus 'ideal' antibiotic prescribing for common conditions in English primary care. *J Antimicrob Chemother*. 2018;73(Suppl 2):19–26.
8. Smieszek T, Pouwels KB, Dolk FCK, Smith DRM, Hopkins S, Sharland M, et al. Potential for reducing inappropriate antibiotic prescribing in English primary care. *J Antimicrob Chemother*. 2018;73(Suppl 2):ii36–43.
9. Pouwels KB, Freeman R, Muller-Pebody B, Rooney G, Henderson KL, Robotham JV, et al. Association between use of different antibiotics and trimethoprim resistance: going beyond the obvious crude association. *J Antimicrob Chemother*. 2018;73:1700–7.
10. Pouwels KB, Muller-Pebody B, Smieszek T, Hopkins S, Robotham JV. Selection and co-selection of antibiotic resistances among *Escherichia coli* by antibiotic use in primary care: an ecological analysis. *BioRxiv*. 2019; 573360. <https://doi.org/10.1101/573360>.
11. Dawson-Hahn EE, Mickan S, Onakpoya I, Roberts N, Kronman M, Butler CC, et al. Short-course versus long-course oral antibiotic treatment for infections treated in outpatient settings: a review of systematic reviews. *Fam Pract*. 2017;34:511–9.
12. Onakpoya IJ, Walker AS, Tan PS, Spencer EA, Gbinigie OA, Cook J, et al. Overview of systematic reviews assessing the evidence for shorter versus longer duration antibiotic treatment for bacterial infections in secondary care. *PLoS One*. 2018;13:e0194858.
13. King LM, Sanchez GV, Bartoces M, Hicks LA, Fleming-Dutra KE. Antibiotic therapy duration in US adults with sinusitis. *JAMA Intern Med*. 2018;178:992–4.
14. Pouwels KB, Hopkins S, Llewelyn M, Walker AS, McNulty CAM, Robotham JV. Antibiotic therapy duration for common infections in English primary care: a cross-sectional analysis and comparison with guidelines. *BMJ*. 2019;364:440.
15. Spellberg B. The new antibiotic mantra – "shorter is better". *JAMA Intern Med*. 2016;176:1254–5.
16. Chai G, Governale L, McMahon AW, Trinidad JP, Staffa J, Murphy D. Trends of outpatient prescription drug utilization in US children, 2002–2010. *Pediatrics*. 2012;130:23–31.
17. Kantor ED, Rehm CD, Haas JS, Chan AT, Giovannucci EL. Trends in prescription drug use among adults in the United States from 1999–2012. *JAMA*. 2015;314:1818–31.
18. Public Health England. Start smart then focus: antimicrobial stewardship toolkit for English hospitals. 2015. <https://www.gov.uk/government/publications/antimicrobial-stewardship-start-smart-then-focus>. Accessed 16 May 2019.
19. Charani E, Castro-Sanchez E, Sevdalis N, Kyratsis Y, Drumright L, Shah N, et al. Understanding the determinants of antimicrobial prescribing within hospitals: the role of "prescribing etiquette". *Clin Infect Dis*. 2013;57:188–96.
20. Public Health England. Management of Infection Guidance for primary Care for Consultation and Local Adaptation. 2017. <https://www.gov.uk/government/publications/managing-common-infections-guidance-for-primary-care>. Accessed 13 July 2018.
21. National Institute for Health and Care Excellence (NICE). Prostatitis (acute): antimicrobial prescribing. Guidance and guidelines. 2018. <https://www.nice.org.uk/guidance/indevelopment/gid-app10007>. Accessed 13 July 2018.
22. Parmar MKB, Carpenter J, Sydes MR. More multiarm randomised trials of superiority are needed. *Lancet*. 2014;384:283–4.
23. Evans SR, Rubin D, Follmann D, Pennello G, Huskins WC, Powers JH, et al. Desirability of outcome ranking (DOOR) and response adjusted for duration of antibiotic risk (RADAR). *Clin Infect Dis*. 2015;61:800–6.
24. Phillips PPJ, Morris TP, Walker AS. DOOR/RADAR: a gateway into the unknown? *Clin Infect Dis*. 2016;62:814–5.
25. O'Riordan W, Tiffany C, Scangarella-Oman N, Perry C, Hossain M, Ashton T, et al. Efficacy, safety, and tolerability of gepotidacin (GSK2140944) in the treatment of patients with suspected or confirmed gram-positive acute bacterial skin and skin structure infections. *Antimicrob Agents Chemother*. 2017;61:e02095–16.
26. Lee JJ, Chu CT. Bayesian clinical trials in action. *Stat Med*. 2012;31:2955–72.
27. Berry SM, Carlin BP, Lee JJ, Muller P. Bayesian adaptive methods for clinical trials. Boca Raton: CRC Press; 2010.
28. Berry DA. Adaptive clinical trials in oncology. *Nat Rev Clin Oncol*. 2012;9:199–207.
29. Quartagno M, Walker AS, Carpenter JR, Phillips PP, Parmar MK. Rethinking non-inferiority: a practical trial design for optimising treatment duration. *Clin Trials*. 2018;15:477–88.
30. van Werkhoven CH, Harbarth S, Bonten MJM. Adaptive designs in clinical trials in critically ill patients: principles, advantages and pitfalls. *Intensive Care Med*. 2019;45(5):678–82. <https://doi.org/10.1007/s00134-018-5426-z>.
31. Curtin F, Heritier S. The role of adaptive trial designs in drug development. *Expert Rev Clin Pharmacol*. 2017;10:727–36.
32. Pallmann P, Bedding AW, Choodari-Oskoei B, Dimairo M, Flight L, Hampson LV, et al. Adaptive designs in clinical trials: why use them, and how to run and report them. *BMC Med*. 2018;16:29.
33. Krams M, Lees KR, Hacke W, Grieve AP, Orgogozo JM, Ford GA, et al. Acute stroke therapy by inhibition of neutrophils (ASTIN): an adaptive dose-response study of UK-279,276 in acute ischemic stroke. *Stroke*. 2003;34:2543–8.
34. Faes C, Aerts M, Geys H, Molenberghs G. Model averaging using fractional polynomials to estimate a safe level of exposure. *Risk Anal*. 2007;27:111–23.
35. Steel MFJ. Model averaging and its use in economics. *arXiv.org*. 2017;arXiv:1709.08221v2. Accessed 13 July 2018.
36. Sydes MR, Parmar MK, Mason MD, Clarke NW, Amos C, Anderson J, et al. Flexible trial design in practice - stopping arms for lack-of-benefit and adding research arms mid-trial in STAMPEDE: a multi-arm multi-stage randomized controlled trial. *Trials*. 2012;13:168.
37. Cellamare M, Ventz S, Baudin E, Mitnick CD, Trippa L. A Bayesian response-adaptive trial in tuberculosis: the endTB trial. *Clin Trials*. 2017;14:17–28.
38. Karrison TG, Huo D, Chappell R. A group sequential, response-adaptive design for randomized clinical trials. *Control Clin Trials*. 2003;24:506–22.
39. Hernán MA, Robins JM. Per-protocol analyses of pragmatic trials. *N Engl J Med*. 2017;377:1391–8.
40. Davies GR, Phillips PPJ, Jaki T. Adaptive clinical trials in tuberculosis: applications, challenges and solutions. *Int J Tuberc Lung Dis*. 2015;19:626–34.
41. Lin J, Bunn V. Comparison of multi-arm multi-stage design and adaptive randomization in platform clinical trials. *Contemp Clin Trials*. 2017;54:48–59.
42. Wason JMS, Trippa L. A comparison of Bayesian adaptive randomization and multi-stage designs for multi-arm clinical trials. *Stat Med*. 2014;33:2206–21.
43. Coffey CS, Levin B, Clark C, Timmerman C, Wittes J, Gilbert P, et al. Overview, hurdles, and future work in adaptive designs: perspectives from a National Institutes of Health-funded workshop. *Clin Trials*. 2012;9:671–80.
44. Bornkamp B, Bretz F, Dmitrienko A, Enas G, Gaydos B, Hsu CH, et al. Innovative approaches for designing and analyzing adaptive dose-ranging trials. *J Biopharm Stat*. 2007;17:965–95.
45. Gillespie D, Francis NA, Carrol ED, Thomas-Jones E, Butler CC, Hood K. Use of co-primary outcomes for trials of antimicrobial stewardship interventions. *Lancet Infect Dis*. 2018;18:595–7.
46. Butler CC, Connor JT, Lewis RJ, Broglio K, Saville BR, Cook J, Van der Velden A, Verheij T. Answering patient-centred questions efficiently: response-adaptive platform trials in primary care. *Br J Gen Pract*. 2018;68:294–5.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

