How much Excess Treatment Failure can a Shorter Antibiotic Course Afford? A Net-Benefit Analysis of Duration Shortening

Dr. Mosrur Ahmed

Assistant Professor, Rahman Institute of Pharmaceutical Sciences and Research, Guwahati, Assam, India.

Corresponding author: laskar.mosrur94@gmail.com

Received: 01 May 2026; Revised: 11 July 2026; Accepted: 17 July 2026; Published: 02 September 2026

Abstract

Background: More than a hundred randomised trials report that shorter antibiotic courses are non-inferior to longer ones, and this has been translated into advocacy for shortening as a default. Non-inferiority establishes the absence of a large disadvantage within a stated margin, and the trade that shortening actually represents — fewer drug-related adverse events against possibly more treatment failures — has not been quantified on a common scale.

Objective: To determine how much excess treatment failure a five-day reduction in therapy can afford before it becomes net harmful, and how much of a conventional non-inferiority margin that permits.

Methods: A decision-analytic model expressed the consequences of shortening on a single scale of harm-equivalents per 1000 patients. Benefit was the adverse events avoided, taken as a rate per antibiotic-day; harm was the excess treatment failures, weighted by w, the number of adverse events considered equivalent to one additional failure. Break-even excess failure was derived analytically. Monte Carlo analysis over the sampling distribution of a trial result gave the probability that shortening is net beneficial for five representative trial outcomes. A policy analysis applied a blanket five-day reduction to a mixture of infections with differing evidence. A threshold analysis determined how large an unmodelled ecological benefit per avoided antibiotic-day would need to be to reverse the conclusion.

Results: At an adverse-event rate of 0.30 % per day and w = 6, shortening by five days avoided 15 adverse events per 1000 patients and broke even at an excess failure of 0.25 percentage points. Across plausible values (adverse-event rate 0.15–0.80 % per day, w = 2–20) the break-even excess ranged from 0.04 to 2.00 percentage points. A conventional 10-point non-inferiority margin was therefore affordable only to the extent of 0.8–7.5 % of its width. The probability that shortening was net beneficial was 89.3 % when the point estimate favoured the short course with a narrow interval, 53.2 % when it was neutral with a wide interval, and 10.5 % when it favoured the long course by 4 points. Applied as a blanket policy to a mixture of infections, net benefit was negative at every weighting examined, ranging from −11.0 to −244.7 harm-equivalents per 1000 patients, with the loss concentrated in infections with weak or contrary evidence.

Conclusions: The case for shortening is strong where the point estimate genuinely favours the short course and weak where non-inferiority rests on a wide margin. Because almost none of a conventional margin is affordable, reporting duration trials as "non-inferior" without the point estimate and margin conceals the quantity on which the decision turns. A blanket default is net harmful in this model unless the ecological benefit per avoided antibiotic-day exceeds 0.0022 to 0.0489 harm-equivalents, a quantity that has never been measured.

Keywords: Antibiotic Duration, Non-Inferiority, Net Benefit, Decision Analysis, Antimicrobial Stewardship, Treatment Failure.

1. INTRODUCTION

The conventional durations of antibiotic therapy were not derived from evidence. Ten days for pneumonia is a convention that outlived its rationale, and early professional guidance conceded in writing that no controlled trial had addressed the question, a concession removed from later editions while the recommendation itself persisted.

Against that convention sits a substantial trial literature. Over twenty-five years more than a hundred randomised controlled trials have compared shorter with longer courses across the common bacterial infections, and their results are consistent: non-inferiority in community-acquired pneumonia, in complicated urinary tract infection, in Gram-negative bacteraemia, in skin and soft tissue infection. Professional guidance has begun to follow, and advocacy has moved from shortening particular durations to shortening as a default.

That move contains a step that is rarely examined. Non-inferiority does not establish equivalence; it establishes that the short course is not worse by more than a pre-specified margin, and those margins are not small. A difference of up to ten percentage points in cure is routinely accepted. Shortening therefore represents a trade rather than a free improvement: fewer adverse drug events and less exposure, against a possible increase in treatment failure bounded only by the margin.

Whether that trade is worth making is a quantitative question and it has not been asked quantitatively. This study asks it. Three questions are addressed: how much excess treatment failure a five-day reduction can afford before the trade turns net harmful; how much of a conventional non-inferiority margin that permits; and what a blanket policy would deliver when applied across infections whose evidence differs.

2. METHODS

2.1 Model structure

Consequences were expressed on a single scale of harm-equivalents per 1000 patients treated. Shortening a course by Δ days avoids adverse drug events at a rate h per antibiotic-day and causes an excess of treatment failures e, expressed in percentage points. Writing w for the number of adverse drug events judged equivalent to one additional treatment failure, net benefit per 1000 patients is

$$NB = 1000 \times (h/100) \times \Delta - w \times 1000 \times (e/100) \quad (1)$$

Setting this to zero gives the break-even excess failure, the largest true increase in failure that shortening can absorb:

$$e_{\text{break-even}} = h\,\Delta / w \quad (2)$$

Equation (2) contains no epidemiology. It is a statement about what the values embedded in the decision imply, and its only inputs are the daily harm rate, the number of days saved and the exchange rate between a failure and an adverse event.

2.2 Parameters

Table 1. Model parameters

Parameter	Symbol	Base case	Range explored	Basis
Days of therapy avoided	Δ	5	—	Typical difference in published duration trials (e.g. 5 vs 10 days)
Adverse drug events per antibiotic-day	h	0.30 %	0.15–0.80 %	Consistent with a reported ≈ 3 % increase in adverse events per additional 10 days
Harm-equivalents per additional treatment failure	w	6	2–20	Stipulated value judgement; varied widely because it is not measurable
Excess treatment failure with the short course	e	varied	0–4 pp	The quantity a non-inferiority trial bounds but does not estimate precisely
Non-inferiority margin	—	10 pp	2.5–10 pp	Margins in common use in duration trials
Monte Carlo draws	—	200,000	—	Sufficient for the reported precision

Note. pp = percentage points. The weighting w is the central value judgement in this analysis and is not an empirical quantity; it is reported across a tenfold range rather than assumed.

The parameter w deserves comment because everything turns on it and it cannot be measured. It expresses how many drug-related adverse events a clinician or health system would accept in order to prevent one additional treatment failure. A value of 2 represents a view that failure is only modestly worse than an adverse event; 20 represents a view that failure is a serious event and adverse drug events are mostly minor. Reasonable people differ, and the results are therefore presented across that range rather than at a preferred value.

2.3 Analyses

Four analyses were performed. Break-even excess failure was computed analytically from equation (2) across the parameter grid. The affordable fraction of a non-inferiority margin was computed as the ratio of the break-even excess to the margin. The probability of net benefit was obtained by Monte Carlo, drawing the true excess failure from the sampling distribution implied by five representative trial outcomes and computing the proportion of draws with positive net benefit. A policy analysis applied a blanket five-day reduction to a mixture of infections weighted by the strength of their evidence, and a threshold analysis determined how large an ecological benefit per avoided antibiotic-day — not otherwise modelled — would be required to reverse the policy conclusion. All computation was performed in Python with a fixed random seed and reproduces exactly.

3. RESULTS

3.1 The trade at base case

Table 2. Net benefit of shortening by five days, base case (h = 0.30 % per day, w = 6)

True excess failure (pp)	Adverse events avoided per 1000	Excess failures per 1000	Net benefit per 1000
0	15.0	0	+15.0
1	15.0	10	−45.0
2	15.0	20	−105.0
3	15.0	30	−165.0
5	15.0	50	−285.0
10	15.0	100	−585.0

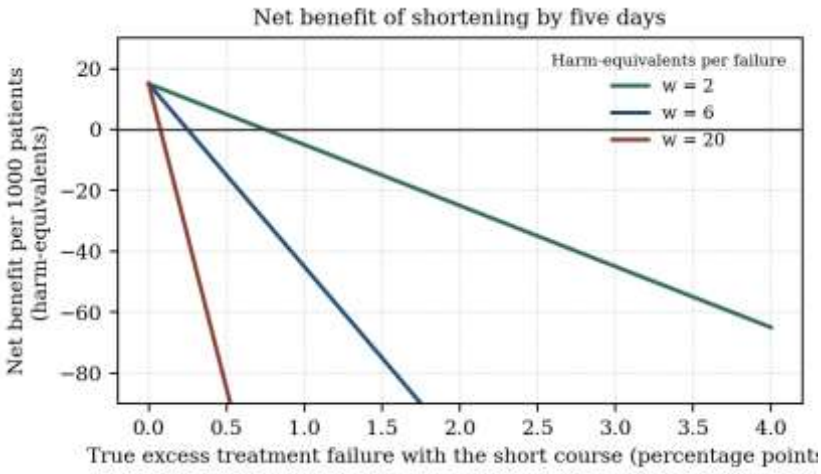

Figure 1. Net benefit of a five-day reduction against the true excess treatment failure, for three weightings.

The asymmetry in Table 2 is the finding. Shortening by five days at an adverse-event rate of 0.30 % per day avoids 15 adverse events per 1000 patients, and that is the entire benefit available. One percentage point of excess failure costs 10 failures per 1000, which at w = 6 is 60 harm-equivalents — four times the whole benefit.

The break-even excess is therefore 0.25 percentage points. Shortening is net beneficial only if the true increase in treatment failure is smaller than one patient in four hundred.

3.2 Break-even across the parameter space

Table 3. Break-even excess treatment failure (percentage points) for a five-day reduction

Adverse events per day	w = 2	w = 4	w = 6	w = 10	w = 20
0.15 %	0.38	0.19	0.12	0.07	0.04
0.30 %	0.75	0.38	0.25	0.15	0.07
0.50 %	1.25	0.62	0.42	0.25	0.12
0.80 %	2.00	1.00	0.67	0.40	0.20

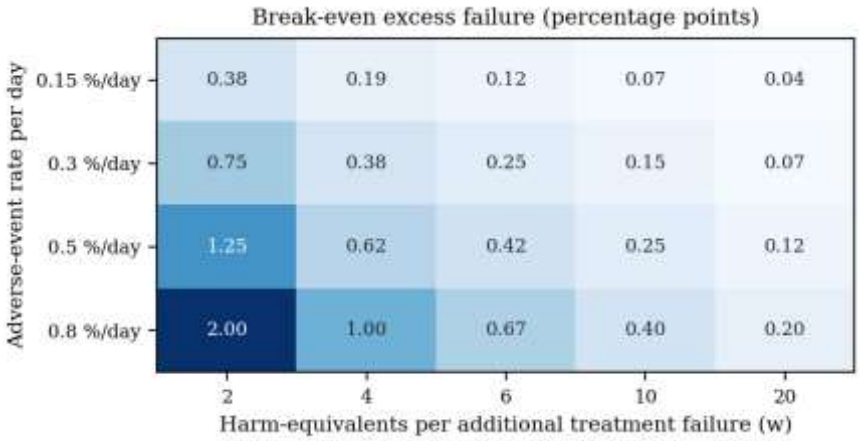

Figure 2. Break-even excess failure across the daily adverse-event rate and the harm weighting.

Across the whole plausible space the break-even excess lies between 0.04 and 2.00 percentage points, and exceeds one percentage point only in the corner where adverse events are frequent and failure is weighted lightly. The conclusion is robust to the value judgement in a specific sense: no defensible choice of w makes a multi-point increase in failure acceptable.

3.3 How much of a non-inferiority margin is affordable

Table 4. Proportion of the non-inferiority margin that is affordable before net harm (h = 0.30 % per day)

Margin	w = 2	w = 4	w = 6	w = 10	w = 20
2.5 pp	30.0 %	15.0 %	10.0 %	6.0 %	3.0 %
5.0 pp	15.0 %	7.5 %	5.0 %	3.0 %	1.5 %
7.5 pp	10.0 %	5.0 %	3.3 %	2.0 %	1.0 %
10.0 pp	7.5 %	3.8 %	2.5 %	1.5 %	0.8 %

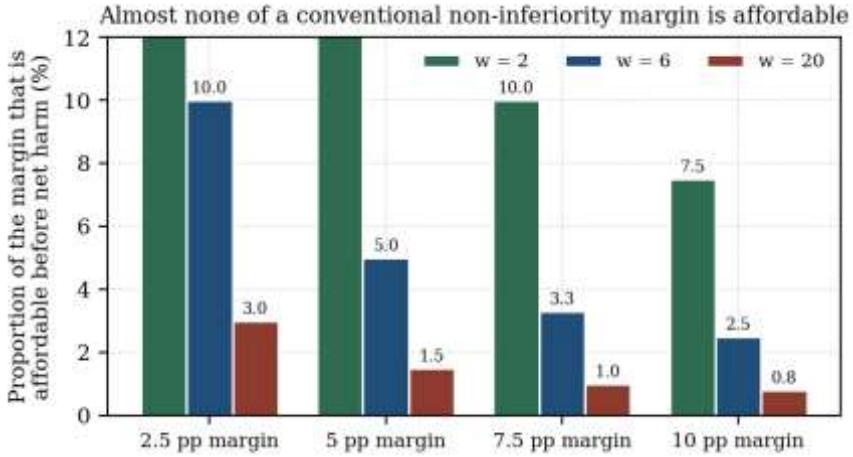

Figure 3. Affordable fraction of conventional non-inferiority margins.

This is the result with the clearest implication for how these trials are read. A trial declaring non-inferiority against a 10-point margin has excluded a true difference larger than 10 points; the analysis here says that only the first 0.8 to 7.5 % of that range is compatible with net benefit. The margin and the decision threshold differ by more than an order of magnitude.

It follows that "non-inferior" is not, on its own, a finding that supports shortening. A trial whose confidence interval sits entirely below zero supports it strongly; a trial whose interval spans zero and reaches nine points supports it hardly at all, and the two are currently reported in the same words. The quantity that matters is the point estimate and the lower part of the interval, not whether the upper bound cleared a margin chosen for feasibility.

3.4 What a given trial result implies

Table 5. Probability that shortening is net beneficial, given the trial result

Trial result (excess failure with short course)	w = 2	w = 6	w = 20
−1.0 pp, standard error 1.0	96.0 %	89.3 %	85.8 %
0.0 pp, standard error 1.0	77.4 %	59.8 %	53.0 %
0.0 pp, standard error 3.0	60.0 %	53.2 %	50.9 %
+2.0 pp, standard error 3.0	34.0 %	28.0 %	26.0 %
+4.0 pp, standard error 3.0	13.8 %	10.5 %	9.6 %

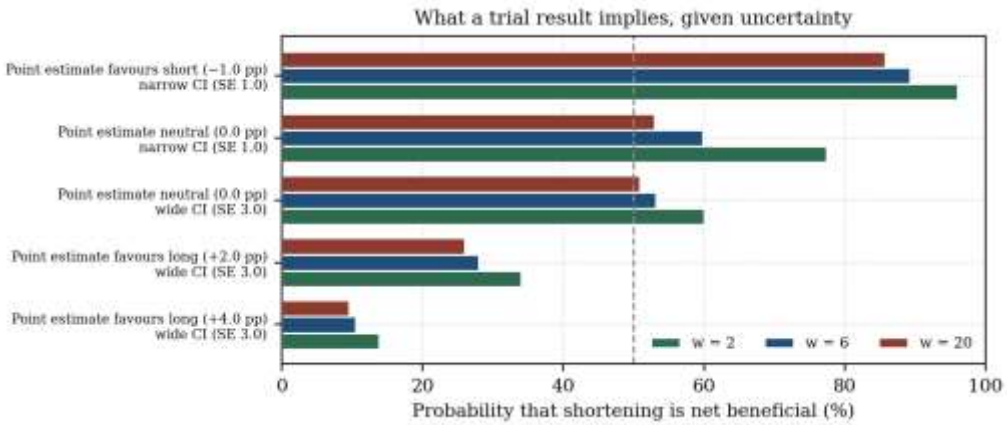

Figure 4. Probability that shortening is net beneficial for five representative trial results.

All five of these results could be reported as demonstrating non-inferiority against a 10-point margin, and they imply very different decisions. A trial with a point estimate favouring the short course and a narrow interval gives an 85.8–96.0 % probability of net benefit and justifies shortening under any weighting. A trial with a neutral point estimate and a wide interval gives 50.9–60.0 %, which is close to a coin toss. A trial whose point estimate favours the long course by four points — still non-inferior at a 10-point margin — gives 9.6–13.8 %, and shortening on that evidence would probably be harmful.

The third row deserves emphasis because it is the commonest situation in this literature: a point estimate near zero with an interval wide enough to satisfy a generous margin. On these numbers such a trial is very nearly uninformative for the decision, and it is exactly the kind of trial that is cited in support of shortening.

3.5 A blanket policy across infections

Table 6. Policy analysis: five-day reduction applied to a mixture of infections

Infection group	Share	Assumed excess failure (pp)	SE	Contribution at w = 6 (per 1000)
Strong short-course evidence	45 %	0.0	1.0	+6.7
Moderate evidence	30 %	1.0	2.0	−13.4
Weak or absent evidence	20 %	3.0	4.0	−33.1
Contrary evidence	5 %	8.0	3.0	−23.3
Total	100 %	—	—	−62.8

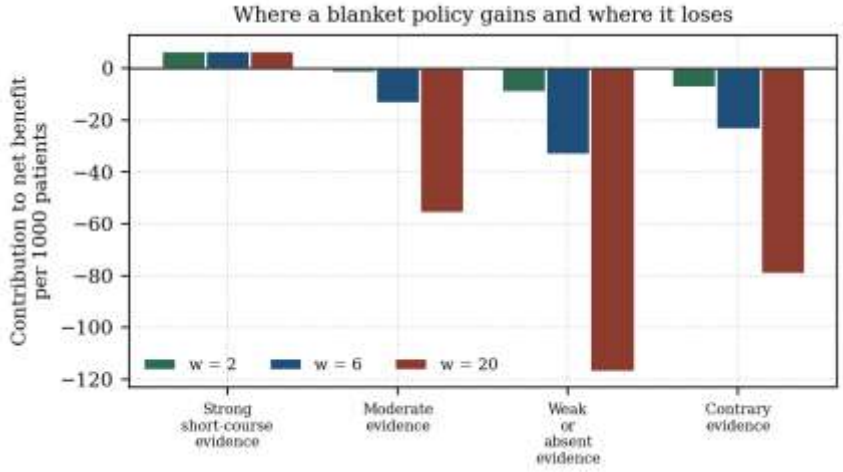

Figure 5. Contribution of each infection group to the net benefit of a blanket policy.

Total net benefit was negative at every weighting examined: −11.0 harm-equivalents per 1000 at w = 2, −62.8 at w = 6 and −244.7 at w = 20. The group with strong evidence contributed positively throughout, at about +6.7 per 1000, and was outweighed by the remainder.

The structure of that result matters more than its sign, which depends on assumptions stated in Table 6 and defended no further than plausibility. A blanket policy earns its benefit in the infections where the evidence is strong and loses it in the infections where the evidence is weak or contrary, and the loss is larger because the excess failure is larger precisely where it has not been excluded. Targeting the policy at the first group would capture the whole of the available benefit and none of the loss, which is an argument for infection-specific shortening rather than for a default.

3.6 What the ecological benefit would have to be

The model above counts only individual-level consequences. The principal argument for shortening is a population one: reduced selection pressure and therefore less resistance. That benefit is real in direction

and has never been quantified per avoided antibiotic-day, so it cannot be entered into the model as a number. It can, however, be solved for.

Table 7. Ecological benefit per avoided antibiotic-day required to make the blanket policy net beneficial

Weighting	Policy net benefit per 1000 (individual level)	Antibiotic-days avoided per 1000	Required benefit per avoided day
w = 2	−11.0	5,000	0.0022 harm-equivalents
w = 6	−62.8	5,000	0.0126 harm-equivalents
w = 20	−244.7	5,000	0.0489 harm-equivalents

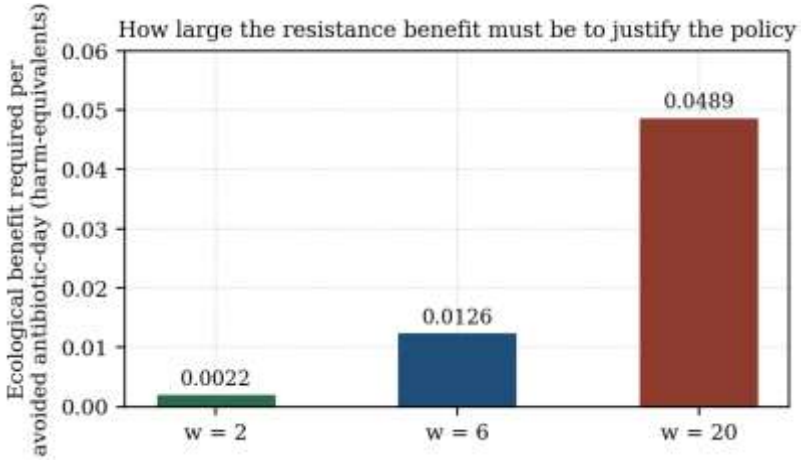

Figure 6. Ecological benefit per avoided antibiotic-day required to reverse the policy conclusion.

The required values are small in absolute terms: at w = 6, an ecological benefit of 0.0126 harm-equivalents per avoided antibiotic-day would make the blanket policy break even, which is roughly one harm-equivalent for every eighty days of therapy avoided. Whether the true value is above or below that is unknown, and this analysis cannot settle it.

This is stated as a genuine limitation rather than a rhetorical concession. If the ecological benefit is of that order or larger, the blanket policy is justified and the individual-level arithmetic in Sections 3.1 to 3.5 is beside the point. The contribution here is to identify the threshold and to observe that a case resting entirely on an unmeasured quantity is being argued as though it rested on the trial evidence, which addresses something else.

4. DISCUSSION

4.1 Principal findings

Shortening therapy by five days buys a small and fixed benefit — 15 adverse events avoided per 1000 patients at a daily event rate of 0.30 % — and can absorb an excess treatment failure of only 0.04 to 2.00 percentage points before turning net harmful, depending on values that reasonable people dispute. Conventional non-inferiority margins are one to two orders of magnitude wider than that break-even. The probability that shortening is net beneficial therefore ranges from 10 % to 96 % across trial results that would all be reported as non-inferior. Applied as a blanket policy across infections of differing evidence, the individual-level net benefit was negative at every weighting, with the loss concentrated where evidence is weakest.

4.2 Implications

Table 8. Implications

Finding	Implication	Recommendation
Break-even excess is 0.04–2.00 pp	The affordable difference is far smaller than the margin	Judge duration trials on the point estimate, not the margin

Only 0.8–7.5 % of a 10-point margin is affordable	"Non-inferior" is not a finding that supports shortening	Require the point estimate and interval in every summary
Probability of net benefit spans 10–96 %	Identical labels conceal opposite decisions	Report probability of net benefit alongside non-inferiority
Neutral estimate with wide interval gives ≈ 50 %	The commonest trial result is nearly uninformative	Do not cite such trials as supporting shortening
Blanket policy loses where evidence is weak	Benefit is concentrated in a minority of infections	Shorten infection by infection, not by default
Required ecological benefit is 0.0022–0.0489 per day	The population case rests on an unmeasured quantity	Measure it, or argue the policy explicitly on that basis
w is not measurable	The decision embeds a value judgement	State the assumed exchange rate when recommending durations

The second row is the most immediately actionable and costs nothing. A duration trial reported as "5 days was non-inferior to 10 days" is compatible with a point estimate favouring either arm and with an interval of any width below the margin. Adding eleven characters — the point estimate and interval — converts a statement that supports nothing in particular into one a prescriber can act on, and the analysis here shows that the difference between those cases is the difference between a 96 % and a 10 % probability of benefit.

4.3 Limitations

The model is deliberately simple and its principal limitation is the one identified in Section 3.6: it counts individual-level consequences and omits the ecological benefit that motivates most stewardship policy. The threshold analysis converts that omission into a stated quantity rather than concealing it, but the conclusion of the policy analysis would reverse if that quantity is large, and nothing here establishes that it is not.

The weighting w is a stipulation and not an empirical parameter. It is reported across a tenfold range and every conclusion is given across that range, but a reader who believes w lies outside 2–20 will reach different numbers. Similarly, treatment failure and adverse drug event are each treated as homogeneous, whereas both range from trivial to fatal; a model distinguishing severity within each category would be more realistic and would require data that does not exist in the duration literature.

The adverse-event rate of 0.30 % per day is derived from a reported increase of approximately 3 % per additional ten days in hospitalised patients, which is a population that differs from the outpatient settings where most duration decisions are made. The parameter is varied across a fivefold range for this reason.

The policy analysis uses an assumed mixture of infections and assumed excess failures for each group. Those values are plausible rather than estimated, and no claim is made that they describe any real case mix; the purpose of that analysis is to show the structure of where a blanket policy gains and loses, not to estimate a national net benefit. Finally, the model is static and ignores that a treatment failure generates a second course of antibiotics, which would add to the exposure the policy was intended to reduce and would make the blanket policy look slightly worse than reported here.

5. CONCLUSION

A five-day reduction in antibiotic therapy avoids about 15 adverse drug events per 1000 patients and can afford an excess treatment failure of a quarter of a percentage point before that benefit is consumed. Conventional non-inferiority margins permit differences forty times larger. The consequence is that the label "non-inferior" carries almost none of the information the decision requires, and trials that would all be described that way imply probabilities of net benefit ranging from one in ten to nineteen in twenty.

Two things follow. Summaries of duration trials should report the point estimate and interval rather than the verdict, because the verdict is determined by a margin chosen for feasibility and the decision is determined by where the estimate actually sits. And shortening should be pursued infection by infection rather than as a default, because in this model the benefit of a blanket policy is concentrated in the minority of infections where the evidence is already strong, while the losses accumulate in those where it is not.

Whether the population-level resistance benefit is large enough to overturn that conclusion is the most important open question in this area, and it is currently assumed rather than measured.

Declarations

Ethical approval: Not applicable.

Funding: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Conflicts of interest: The authors declare that there are no conflicts of interest regarding the publication of this paper.

Data availability: The data that support the findings of this study are available from the corresponding author upon reasonable request.

Author contributions:

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Dr. Mosrur Ahmed		✓		✓		✓	✓	✓		✓		✓		✓

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

6. REFERENCES

1. B. Spellberg, "The new antibiotic mantra: shorter is better," JAMA Intern Med, vol. 176, no. 9, pp. 1254-1255, 2016. doi.org/10.1001/jamainternmed.2016.3646
2. R. A. Lee, R. M. Centor, and L. L. Humphrey, "Appropriate use of short-course antibiotics in common infections: best practice advice from the American College of Physicians," Ann Intern Med, vol. 174, no. 6, pp. 822-827, 2021. doi.org/10.7326/M20-7355
3. L. B. Rice, "The Maxwell Finland Lecture: for the duration - rational antibiotic administration in an era of antimicrobial resistance and Clostridium difficile," Clin Infect Dis, vol. 46, no. 4, pp. 491-496, 2008. doi.org/10.1086/526535
4. S. Royer, K. M. Demerle, R. P. Dickson, and H. C. Prescott, "Shorter versus longer courses of antibiotics for infection in hospitalized patients: a systematic review and meta-analysis," J Hosp Med, vol. 13, no. 5, pp. 336-342, 2018. doi.org/10.12788/jhm.2905
5. D. Yahav, E. Franceschini, and F. Koppel, "Seven versus 14 days of antibiotic therapy for uncomplicated Gram-negative bacteremia: a noninferiority randomized controlled trial," Clin Infect Dis, vol. 69, no. 7, pp. 1091-1098, 2019. doi.org/10.1093/cid/ciy1054
6. Y. Mo, S. Booraphun, and A. Y. Li, "Individualised, short-course antibiotic treatment versus usual long-course treatment for ventilator-associated pneumonia (REGARD-VAP): a multicentre, individually randomised, open-label, non-inferiority trial," Lancet Respir Med, vol. 12, no. 5, pp. 399-408, 2024. doi.org/10.2139/ssrn.4492691
7. A. Hoberman, J. L. Paradise, and H. E. Rockette, "Shortened antimicrobial treatment for acute otitis media in young children," N Engl J Med, vol. 375, no. 25, pp. 2446-2456, 2016. doi.org/10.1056/NEJMoa1606043
8. D. J. Williams, C. B. Creech, and E. B. Walter, "Short- vs standard-course outpatient antibiotic therapy for community-acquired pneumonia in children: the SCOUT-CAP randomized clinical trial," JAMA Pediatr, vol. 176, no. 3, pp. 253-261, 2022. doi.org/10.1001/jamapediatrics.2021.5547
9. P. D. Tamma, E. Avdic, D. X. Li, K. Dzintars, and S. E. Cosgrove, "Association of adverse events with antibiotic use in hospitalized patients," JAMA Intern Med, vol. 177, no. 9, pp. 1308-1315, 2017. doi.org/10.1001/jamainternmed.2017.1938

10. J. Curran, J. Lo, and V. Leung, "Estimating daily antibiotic harms: an umbrella review with individual study meta-analysis," Clin Microbiol Infect, vol. 28, no. 4, pp. 479-490, 2022. doi.org/10.1016/j.cmi.2021.10.022
11. G. Piaggio, D. R. Elbourne, S. J. Pocock, S. Evans, and D. C. Altman, "Reporting of noninferiority and equivalence randomized trials: extension of the CONSORT 2010 statement," JAMA, vol. 308, no. 24, pp. 2594-2604, 2012. doi.org/10.1001/jama.2012.87802
12. T. R. Fleming, "Current issues in non-inferiority trials," Stat Med, vol. 27, no. 3, pp. 317-332, 2008. doi.org/10.1002/sim.2855
13. L. Mauri and D. Agostino, "Challenges in the design and interpretation of noninferiority trials," N Engl J Med, vol. 377, no. 14, pp. 1357-1367, 2017. doi.org/10.1056/NEJMra1510063
14. A. J. Vickers and E. B. Elkin, "Decision curve analysis: a novel method for evaluating prediction models," Med Decis Making, vol. 26, no. 6, pp. 565-574, 2006. doi.org/10.1177/0272989X06295361
15. A. H. Briggs, K. Claxton, and M. J. Sculpher, Decision Modelling for Health Economic Evaluation. Oxford: Oxford University Press, 2006. doi.org/10.1093/oso/9780198526629.001.0001
16. M. J. Llewelyn and J. M. Fitzpatrick, "The antibiotic course has had its day," BMJ, vol. 358, 2017. doi.org/10.1136/bmj.j3418
17. A. C. Kalil, M. L. Metersky, and M. Klompas, "Management of adults with hospital-acquired and ventilator-associated pneumonia: 2016 clinical practice guidelines by the IDSA and ATS," Clin Infect Dis, vol. 63, no. 5, pp. e61-e111, 2016. doi.org/10.1093/cid/ciw353