BMJ Open

Point of Care Tests, Diagnostic Uncertainty and Antimicrobial Stewardship in the ICU: Procalcitonin or PCR to aid antibiotic stop decisions

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2024-084872
Article Type:	Original research
Date Submitted by the Author:	04-Feb-2024
Complete List of Authors:	Lau, Timothy; Imperial College London, Nurek, Martine; Imperial College London, Surgery and Cancer Singhal, Archit; Imperial College London Moore, Luke; Chelsea and Westminster Hospital NHS Foundation Trust, Department of Microbiology; Imperial College London, National Institute for Health Research Health Protection Research Unit in Healthcare Associated Infections and Antimicrobial Resistance Mughal, Nabeela; Imperial College London; Chelsea and Westminster Hospital NHS Foundation Trust Singh, Suveer; Imperial College London, Faculty of Medicine; Chelsea and Westminster Hospital NHS Foundation Trust, Respiratory and Intensive Care Medicine
Keywords:	Behavior, MICROBIOLOGY, INFECTIOUS DISEASES, INTENSIVE & CRITICAL CARE, Diagnostic microbiology < INFECTIOUS DISEASES

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Point of Care Tests, Diagnostic Uncertainty and Antimicrobial Stewardship in the ICU: Procalcitonin or PCR to aid antibiotic stop decisions

Lau, Timothy[1,§], Nurek, Martine[1,§], Singhal, Archit[1], Moore, Luke[1,2], Mughal, Nabeela[1,2], Singh, Suveer[1,2,§,*]

[1] Imperial College London, London, UK

[2] Chelsea and Westminster Hospital NHS Foundation Trust, London, UK

[§] These authors contributed equally to this work

*Correspondence to:

Professor Suveer Singh

Departments of Adult Intensive Care Medicine and Respiratory Medicine, APMIC, Imperial College London, 369 Fulham Road, London SW10 9NH, UK

suveer.singh@imperial.ac.uk

WORD COUNT: 4766

ABSTRACT

OBJECTIVES: ICU clinicians stop antibiotics more often, by a negative infection-point of care test (PCR-POCT). Simulated cases of diagnostic uncertainty regarding infection resolution, led clinicians to choose options and aid stop decisions; Procalcitonin (PCT) and/or PCR-POCT tests, de/escalation. We hypothesized that a direct infection indicator, PCR-POCT would influence stop judgements moreso than indirect PCT. Accordingly, we tested antibiotic stop decisions when presented a negative PCR-POCT, despite borderline positive PCT.

DESIGNS: Observational prospective study

SETTING: Intensive Care Unit

PARTICIPANTS: 66 ICU clinicians in University hospitals.

METHODS: Clinicians saw 4 scenarios of different clinico-biological trajectories:1) clear improvement, 2) clear worsening, 3) discordant-clinically better/biologically worse, 4) discordant-clinically worse/biologically better. Participants gave an initial decision (stop/ continue/continue-escalate/continue-de-escalate). Then PCR-POCT and/or PCT offered (accept/decline). Irrespective, after a negative PCR-POCT and borderline positive PCT result, a final antibiotic decision was made.

MEASURES: Proportion of stop decisions before vs. after test results per scenario. The association of the final decision with the clinician's change in confidence, willingness to request the biomarker(s), and the case trajectory were determined.

RESULTS: Fewer clinicians than expected stopped antibiotics v baseline (36%, 94/264 vs 42%, 110/264, p=0.045). This was so in 3 of 4 scenarios, significantly less in the improvement (p<0.001) and the discordant clin-better scenario (p=0.024). PCT was requested more frequently than PCR-POCT (61% vs. 53%, p<0.001). PCT requesters (v declining) were significantly less inclined to stop antibiotics (p<0.001),

whilst PCR-POCT requesting led to more stopping (p<0.001), before knowing test results.

CONCLUSIONS: A negative PCR-POCT result did not increase clinicians' inclination to stop antibiotics, when alongside a borderline positive PCT. This reflects clinicians' natural risk aversion. PCT was more popular than PCR-POCT, but PCR-POCT was more likely to aid stop decisions.

Their comparison, role, utility and selective deployment for influencing antibiotic stop decisions more effectively requires a large RCT.

STRENGTHS AND LIMITATIONS OF THIS STUDY

- In four typical clinical vignettes, with different clinico-biological trajectories, this study offered a realistic simulation of ICU-related respiratory infection factors to test antibiotic stewardship decisions.
- Choices to escalate/de-escalate antibiotics alongside stop and continue options, provide a reproducible and adaptable test platform to study the clinical situational and behavioural factors that influence antibiotic decisions.
- By focussing on the end (rather than onset) stage of infection, it offered the opportunity to test clinicians' preferences for direct (PCR-POCT) or indirect (PCT) point of care tests as arbiters for antibiotic decisions, when confronted with clinic-biological diagnostic uncertainty.
- A larger sample is required to robustly determine preferences between PCR-POCT or PCT, and influential factors. This would have to utilize the same choices (e.g. continue/stop/(de)escalate), whilst adding the presentation of theoretical combinations of positive and negative results, using the same test vignette platform.

BACKGROUND

Antimicrobial resistance has become an increasingly pertinent issue within patient care (1). Antibiotic stewardship programmes (ASPs) are strategies that aim to improve the use of antibiotics and have been employed to reduce the potential for antimicrobial resistance. Antibiotic prescribing is high in the ICU setting, with up to 70% of patients having an antibiotic prescribed and so the use of ASPs in this setting can have large effects on antimicrobial resistance (2). Some ASPs have had success in reducing antibiotic usage, including procalcitonin (PCT)-guided antibiotic stewardship for sepsis in the ICU (3). PCT is a surrogate biomarker for infection and is considered a useful tool in antibiotic prescribing, having been successful in reducing antibiotic course lengths in study settings (4-7). Another useful tool to potentially improve antibiotic prescribing is the polymerase chain reaction point of care test (PCR-POCT, herein also referred to as POCT). These tests have been used during the COVID-19 pandemic and are found to have high diagnostic accuracy (8). Infection identifying POCTs can effectively rule out the presence of infective organisms; thereby increasing clinicians' confidence to stop antibiotics (9). This is important as clinicians are found to continue (rather than stop) empirical antibiotics when there is clinical uncertainty, based on natural risk averseness, particularly with more severe illness (10). POCT can alleviate this uncertainty and should therefore reduce antibiotic prescribing. However recent studies have not seen this. In the VAPrapid trial of ventilator associated pneumonia, use of a highly accurate cytokine based POCT (interleukin 1/8) rule out test failed to increase antibiotic-free days in ICU despite excellent test performance (11).

One possible explanation is that factors such as cognitive biases, may be impacting the use of these tests and antibiotic decision making more broadly. Confirmation

bias, whereby clinicians find and interpret evidence to support an existing hypothesis, could affect decision making (12,13). Anchoring, whereby clinicians fixate on early information in a case, may lead clinicians to overlook important possibilities (14). While these clinician factors have been acknowledged in previous work, research investigating their specific and quantifiable effect on decision making is lacking (15).

A recent study investigated clinical and clinician factors that influence POCT-use and antibiotic prescribing in ICU (16). This vignette-based experiment found that a negative PCR-POCT result (suggesting no infection) significantly increased clinicians' inclination to stop antibiotics, but three "competing" factors worked to decrease it: an ambiguous or worsening patient trajectory, clinicians' first impressions (i.e., high confidence that antibiotics were needed) and lack of interest in POCT (i.e., rejection of the POCT when it was offered). Whilst that study highlighted the potential utility of POCT for antibiotic cessation in the setting of ICU related respiratory infection, as well as the factors that might diminish its utility, the model was simple with limitations. Four vignettes describing patients that had just completed a course of antibiotics were constructed: one was clearly improving (clinical and biological signs better than upon admission), one was clearly deteriorating (clinical and biological signs worse), and two were ambiguous (clinical signs better but biological signs worse / clinical signs worse but biological signs better). The study found that POCT was most requested and most effective (in promoting antibiotic stop decisions) in the ambiguous scenarios that featured clinical worsening but biological improvement (and overall worsening). This is important, as a critical component of increasing and streamlining POCT use in the ICU is identifying scenarios in which it is most/least helpful. For reliability, those findings

would benefit from replication in a new set of vignettes that adhere to the same structure (i.e., a new vignette depicting clear improvement, a new vignette depicting clear worsening, etc).

In the original study, participants' uptake of the POCT offer was relatively high (overall 65% of the time). In clinical practice, however, clinicians have numerous tests available, many of which are more established and therefore more widely used than POCT (e.g., procalcitonin, PCT). Whether clinicians would still request POCT when PCT is available (albeit an indirect biomarker surrogate of an infectious agent rather than a direct form in PCR-POCT) remains to be seen.

Thirdly, in the original study, clinicians were only asked to choose between two courses of action: stop or continue antibiotic treatment, in reality, the opportunity to (de)escalate, are further options. Recognising these limitations, the authors asked participants at the end of the study: "*Had the option to de/escalate antibiotics been available, would you have used it?*" Most clinicians (74.3%, 52/70) responded yes, leading the authors to conclude that findings may have looked different had (de)escalation been available.

The present study aimed to address these limitations. Specifically, to replicate and validate the previous findings, by making three improvements to the study design:

1. clinicians were presented with a new set of similar vignettes (one improving, one worsening, two ambiguous) assessing reliability of the model;
2. clinicians were offered a POCT and a PCT test (they could select either, neither, or both) with the two providing conflicting results (POCT negative, PCT positive). This allowed us to assess whether there is a systematic

preference for either POCT or PCT, and how clinicians consolidate disparity between the two in their decision making;

3. clinicians had the option to stop, continue, escalate or de-escalate antibiotics.

We hypothesised that a negative POCT result would increase clinicians' inclination to stop antibiotics, as in the previous study (*hypothesis 1*) (16). We also expected that this effect would be smaller than that observed in the previous study (*hypothesis 2*), because:

1. the presence of a borderline positive PCT result would reduce clinicians' inclination to stop, because clinicians may have greater trust in the (more established) PCT test than the (less established) PCR-POCT, or because its implication of the possible presence of infection may mitigate the POCT result;
2. allowing clinicians to (de)escalate antibiotics would reduce the incidence of stopping, as de-escalation may be perceived as a "safer" (i.e., less risky) alternative.

METHODS

Participants

Consultants and trainees from the ICU (with 3+ months continuous experience in ICU) currently working in London-based university hospitals were invited through advertisements in closed social media groups exclusive to ICU consultants and trainees. Clinicians who took part in the previous study (16) and had indicated a willingness to be involved in future studies, were contacted via direct email. The

email contained a direct link to the online survey hosted by Qualtrics (Washington, USA). The survey remained open from March – October 2022.

Sample size

A minimum of 77 responses were required to demonstrate the same effect size of an increase in antibiotic stop decisions from 54-70% (16) with 80% power at the alpha = 0.05 level. Further details are available in the Supplemental Materials (SM1).

Importantly, however, we expected the effect to be smaller in the present study than the previous one (see *hypothesis 2*). We therefore conducted a second sample size calculation, identical to the first except that it aimed to detect a smaller effect (w =0.20). After adjusting for clustered data, the number of responses required was 233 and the number of participants required was 58 (233/4). We therefore aimed to recruit 58 participants.

Materials

We constructed four clinical scenarios of resolving lung infection after a course of antibiotics. Each scenario comprised clinical and biological data, which were varied to create four distinct patient trajectories (Table 1).

Table 1. Clinical vignettes used in this study. Full details of the scenarios can be found in the Supplemental Materials (SM2).

Vignette name	**Description**
Improvement	A 68-year-old man with lobar pneumonia. Clinico-biological improvement after a 5-day course of antibiotics.

Worsening	A 78-year-old man with lobar pneumonia who deteriorates to requiring mechanical ventilation. After initial stabilisation after 5 days of antibiotics, there is a decline in clinical and biological status.
Discordant: clinically better, labs worse ("disc clin better")	A 60-year-old man with a severe lobar pneumonia requiring mechanical ventilation who improves clinically at 7 days and is extubated after an antibiotic course, but whose blood biomarkers are worse.
Discordant: clinically worse, labs better ("disc clin worse")	A 59-year-old man with multilobar pneumonia who completes a course of antibiotics, is extubated but then deteriorates clinically despite improving blood biomarkers of infection.

These vignettes were thought to accurately represent the varying degrees of diagnostic (un)certainty commonly encountered in the critical care setting. Two of the scenarios (*improvement* and *worsening*) acted as 'positive' and 'negative' controls, in that the trajectory clearly supported stopping antibiotics (*improvement*) or continuation/escalation of antibiotics (*worsening*). The remaining two scenarios presented uncertainty with regards to antibiotic decision making and were used to explore the importance of the clinical vs. biological trajectories.

In each scenario, two test results were made available: POCT and PCT. Both tests were described as valid and reliable. The POCT was an infection-identifying PCR test that provided rapid diagnostics for bacteria. The POCT result (unbeknown to clinicians at the point of request) was always negative (suggesting no active infection within the lung), while the PCT result was always positive (suggesting infection). Notably, the POCT in these scenarios was an indicator of the actual presence of an infectious organism, whereas the PCT test was a surrogate marker for the presence of infection. The conflicting results of these two tests would allow us to compare how a direct vs. indirect test for infection might influence clinicians' choices.

Procedure

Clinicians provided informed consent. Following this, they responded to the four vignettes, presented in a random order. Each vignette began with a brief patient description, including the patient's age, sex, details of admission and clinical status one week later (i.e., after completing a course of antibiotics). Based on this, clinicians were asked to decide the best course of action with regards to antibiotics (*stop antibiotics* / *continue with current course of antibiotics* / *escalate antibiotics* / *de-escalate antibiotics*). They were also asked to rate their confidence in this decision, on a 6-point Likert scale anchored at 1=*not at all confident* and 6=*extremely confident*.

Clinicians were then informed of the availability of a POCT and a PCT test. Clinicians could select the test/s that they wished to perform (*POCT only / PCT only / both POCT and PCT / neither POCT nor PCT*). They were also asked to indicate the reason(s) for this decision (SM3).

Clinicians that chose to perform one of the two tests (*POCT only / PCT only*) were at this point presented with the result of the chosen test (POCT was always negative, PCT always positive) and asked whether they would change their previous antibiotics decision (*yes / no*). Those who responded *yes* were offered the previous antibiotics options again (*stop / original course / escalate (more than original course) / de-escalate (less than original course)*). Regardless of whether their decision had changed, they were asked to indicate their confidence in their decision (1-6, as above) and to explain their decision (free text response). Following this, the second (non-requested) test result was displayed (POCT was always negative / PCT was always positive) and clinicians were asked the same questions again; that is, they were asked whether they would change their decision (if *yes*, they were presented

with the four antibiotics options described above), to indicate their confidence in their decision (1-6), and to explain their decision (free text).

Clinicians that chose to perform both or neither of the tests (*both POCT and PCT / neither POCT nor PCT)* were shown the results of both tests simultaneously (POCT negative, PCT positive) and asked exactly the same questions; i.e., whether they would change their decision (*yes / no*) and what they would change it to if so (*stop / original course / escalate (more than original course) / de-escalate (less than original course)*), confidence in decision (1-6) and rationale for decision (free text).

Importantly, therefore, clinicians always gave an "initial" antibiotic decision (no test results seen) and a "final" antibiotic decision (both test results seen). When clinicians chose to perform one of the two tests, they also gave "interim" decisions (that one test result seen).

After completing all four scenarios, clinicians were also asked to complete Grol et al.'s *Attitudes to Risk-Taking in Medical Decision Making* (17) questionnaire (adapted to the ICU setting) and to provide demographic information (gender and level of training). The study procedure is presented graphically in the Supplemental Materials (SM4). Details of the piloting processes are presented in the Supplemental Materials (SM5).

Statistical analysis

To measure the combined effect of the test results (negative POCT and positive PCT) on clinicians' inclination to stop antibiotics, we compared the proportion of clinicians that chose to stop antibiotics initially vs. finally, using chi-square analysis. We had intended to also measure and compare the effects of each test result by

analysing the interim decisions of those who selected *POCT only* or *procalcitonin only*; however, very few participants selected these options (see Results), rendering such an analysis unreliable.

To explore this further, we created a continuous variable termed "willingness-to-stop", by signing confidence ratings in accord with antibiotic decisions. Specifically, confidence ratings (1=*not at all confident* to 6=*extremely confident*) were signed positive (+) if clinicians chose to stop antibiotics, or negative (-) if they chose to continue (be it in the form of *escalation*, *de-escalation*, or *continuation of original course*). Initial and final willingness-to-stop antibiotics (both ranging from -6=lowest to 6=highest) were compared using Wilcoxon signed ranked tests (non-parametric, as the willingness-to-stop variables were not normally distributed, see Results).

We also explored whether clinicians might have a preference for POCT vs. PCT tests, by examining the proportion of participants that requested (vs. rejected) each test. We explored whether this might differ by scenario, using chi-square analysis.

Finally, we explored the effect of scenario (1=*improvement*, 2=*disc clin better/worse*, 3=*worsening*), initial antibiotic decision (0=*escalate*, 1=*original course*, 2=*de-escalate*, 3=*stop*), test(s) requested (1=*requested* and 0=*rejected* for POCT and procalcitonin, respectively), *Attitude To Risk-Taking* score (the per-participant sum of responses to Grol et al.'s questionnaire), and level of experience (0=trainee, 1=consultant) on final antibiotic decisions (0=*escalate*, 1=*original course*, 2=*de-escalate*, 3=*stop*), using a mixed effects linear regression model with a per-participant random intercept. While this model does not require a normally distributed dependent variable, we repeated it using an ordinal (rather than linear) procedure, to see whether findings changed.

A *p*-value < 0.05 was considered as statistically significant. Statistical analysis and graphing were performed using SPSS 28 (IBM, New York, USA) and Stata/MP 17 software (StataCorp, Texas, USA).

Approval

This study was approved by the Imperial College Research Ethics Committee (ICREC reference 20IC6499) and the manuscript adheres to CHERRIES guidance for reporting e-survey results (18).

Patient and public involvement

None.

RESULTS

Demographic data

Sixty six clinicians completed the survey. The number of clinicians that accessed the survey is unknown, as incomplete responses were deleted automatically, following one week of inactivity. All participants completed all four scenarios, providing a total of 264 scenario responses. Of the 66 clinicians, 39 (59.1%) were male. There were 34 (51.5%) consultants, 17 (25.8%) SpR clinicians, 13 (19.7%) SHO clinicians and 2 (3%) Foundation Year clinicians. The demographics of the sample can be found in the Supplemental Materials (SM6).

Initial and final antibiotic decisions

Figure 1 shows participants' initial and final antibiotic decisions. Prior to receiving any test results (POCT or PCT), participants opted to stop antibiotics 41.7% of the time (110/264). This was reduced to 35.6% (94/264) following the negative POCT and positive PCT test. Therefore, clinicians were less likely (rather than more likely) to stop antibiotics after receiving the test results (chi-square=4.03, df =1, p=0.045). Few clinicians chose the path of de-escalation initially or subsequent to the POCT/PCR results, with most opting to stop or escalate.

The frequency of stop decisions within each scenario is displayed in Figure 2. Within the *improvement*, *disc clin better* and *disc clin worse* scenarios, fewer participants chose to stop antibiotics in their final decision (following both test results) compared to their initial decision. This effect was highly significant in the *improvement* scenario (72.7% vs. 86.4%; chi square=10.50, df=1, p=0.001), significant in *disc clin better* (28.8% vs. 42.4%; chi-square=5.01, df=1, p=0.025), and non-significant in *disc clin worse* (36.4% vs. 37.9%, chi-square=0.07, df=1, p=0.797). In the *worsening* scenario, no participants chose to stop antibiotics initially, and was statistically unchanged following the test results (4.5%, p=*ns*). For a full breakdown of initial and final antibiotic decisions per scenario (i.e., number of participants that elected to *stop*, *continue with the original course*, *escalate*, and *de-escalate*), see the Supplemental Material (SM7).

Initial and final willingness-to-stop were nonparametric following inspection of frequency distribution histograms. This was confirmed using Kolmogorov-Smirnov tests (p<0.001 for both initial and final willingness-to-stop). Figure 3 shows clinicians' median willingness-to-stop before and after receiving the negative POCT and positive PCT results, per scenario. Similar to Figure 2, willingness-to-stop appeared

to decrease from initial to final decision in the *improvement*, *disc clin better* and *disc clin worse* scenarios; the effect was significant in the *improvement* scenario (z=-3.84, p<0.001, effect size r=0.33), significant in the *disc clin better* scenario (z=-2.56, p=0.010, r=0.22), and non-significant in the *disc clin worse* scenario (z=-0.11, p=0.909, r=0.01). In the *worsening* scenario, willingness-to-stop was greater in the final decision than the initial decision, suggesting that the negative POCT might increase clinicians' willingness-to-stop in this specific scenario. This effect was significant (z=-3.04, p=0.002, r=0.26). The evolution of the willingness to stop from the initial decision, through an interim option and then the final decision is shown in the Supplemental Material (SM8).

POCT and PCT requests

Across all scenarios, POCT was requested 53.4% of the time (141/264), however this varied significantly by scenario (chi-square=55.97, df=3, p<0.001; *improvement=15.2%, disc clin better=57.6%, disc clin worse=68.2%, worsening=72.7%*). Similarly, PCT was requested 61% of the time (161/264) and also varied significantly by scenario (chi-square=52.01, df=3, p<0.001; *improvement=24.2%, disc clin better=74.2%, disc clin worse=78.8%, worsening=66.7%*).

Figure 4 displays the tests requested by clinicians per scenario. In the *improvement* scenario, the majority of participants requested neither the POCT nor PCT test (72.7%), potentially due to the patient's unambiguously positive trajectory (ceiling effect). Within the *disc clin better, disc clin worse* and *worsening* scenarios, most participants requested both tests (53.0%, 60.6% and 60.6% respectively).

Participants' reasons for requesting/rejecting the tests were mainly related to supplementing clinical judgement and deeming tests necessary or not (SM9).

Factors influencing the final decision

Our mixed effects model explored the effect of 1) initial antibiotic decisions, 2) test/s requested, 3) attitudes toward risk taking, 4) levels of experience, and 5) scenarios on final antibiotic decisions. The results showed that initial antibiotic decisions had a strong effect upon final antibiotic decisions (b=0.78 [95% confidence interval 0.68 to 0.87], p<0.001). That is, clinicians who were more [less] inclined to stop antibiotics prior to receiving the test results were also more [less] inclined to stop after receiving them. This was consistent with the original study (16). Additionally, participants that requested PCT (either alone or in conjunction with POCT) were less inclined to stop in their final antibiotics decision (b=-0.48 [-0.69 to -0.26], p<0.001), whereas participants that requested POCT (either alone or in conjunction with PCT) were more inclined to stop in their final antibiotics decision (b=0.41 [0.19 to 0.64], p<0.001). This was consistent with the original study (16). Level of experience (trainee or consultant) had no significant effect on final antibiotics decisions (b=0.14 [-0.06 to 0.33], p=0.158), nor did differences in attitudes towards risk-taking (b=0.01 [-0.02 to 0.04], p=0.572) or scenario (b=-0.14 [-0.32 to 0.03], p=0.102). These findings did not change when we reran the analysis, using an ordinal rather than a linear model (SM10), nor when we replaced initial and final antibiotic decisions with initial and final willingness-to-stop (SM11).

DISCUSSION

Antibiotic stop decisions did not increase following a negative PCR POCT result and borderline positive PCT result, after a completion course of antibiotics in ICU related respiratory infection. Rather, in three out of four scenarios (*improvement*, *disc clin better* and *disc clin worse*), the stop rate decreased. This differed from the first WHYSTOP study (16), where antibiotic stop decisions increased consistently (i.e., in all scenarios) following receipt of a negative POCT. Thus, a negative POCT did not increase stop decisions when there was a borderline positive PCT result.

A few potential explanations for the lack of increase in antibiotic stop decisions are likely. First, the addition of a borderline positive PCT result likely negated the effect of the negative POCT and so reduced clinicians' inclination to stop. Second, participants may have been more inclined to trust the PCT result (vs newer non established PCR POCT) due to its known potential and growing evidence base within clinical medicine (4,6,7). Alternatively, participants may have had equal trust in PCT and POCT, but conflicting test results triggered a conservative approach to 'err on the side of caution' (i.e., continue antibiotics) (10). This was likely, as demonstrated in the original WHYSTOP study. Specifically, in that study, the improvement scenario was extended to include the following 'twist': the original negative PCR POCT result was declared erroneous (due to lab error) and re-testing gave a positive result. In light of this, clinicians were given the opportunity to revise their antibiotic decisions. The proportion of antibiotic stop decisions, even in this clinico-biological case of improvement (suggesting the resolution of infection), fell from 90% to 61% (16). The caution adopted relates to 'prospect theory'; the idea that a loss and regret is psychologically twice as impactful as a gain; the context being potential patient harm from an antibiotic stop decision (19,20). Another explanation may be that the clin better, biologically worse scenario may have been slightly more ambiguous in its clinical trajectory than in the

original paper – that may have accounted for less confidence in choosing to stop antibiotics post POCT.

Finally, we expected that the addition of a de-escalation option might reduce the incidence of stopping antibiotics (by providing a “less risky” alternative) and thus weaken the effect of the negative POCT. Contrary to expectation, the de-escalation option was rarely used (8.1% of initial decisions and 9.4% of final decisions) and is therefore unlikely to account for the present findings.

Participants who actively requested (vs. passively received) the POCT result were more inclined to stop antibiotics, whereas those who actively requested (vs. passively received) the PCT result were less inclined to stop. Clearly, interpretation and/or weighting of test results does not take place in isolation, but is dependent upon the perceived relevance, and/or value of the test in the decision making. This suggests that ‘forcing’ POCT on clinicians is unlikely to bring about effective change. Rather increasing awareness and trust in POCT may allow for greater confidence in its use and increase uptake. Indeed, Dhesi et al. identified lack of clinician trust as a potential barrier to POCT adoption in practice, which must be addressed if POCT is to be of value in ICU settings (21). This was also a potential reason for the lack of improvement in patient outcomes in a quasi-randomised controlled trial, where use of PCR POCT was compared with routine laboratory PCR in guiding clinical management (22).

Clinicians’ grade (consultant vs. trainee) did not influence the inclination to stop antibiotics, nor did clinicians’ attitude towards risk taking. The WHYSTOP study was underpowered to address this (16). However, in a subsequent study, a difference in clinical decision making was noted between novices (i.e. clinical medical students) and

clinicians (23). This may be of particular interest, given the potential for learnt behaviour on antibiotic decision making (24,25). Indeed, medical students, tested in the same conditions of the original WHYSTOP study, were initially more conservative than ICU clinicians in STOP decisions. More students chose the POCT when offered than clinicians, and the negative result increased the proportion of stop decisions in all scenarios, to the same level as clinicians (23).

We made changes to the WHYSTOP study protocol to better reflect reality. These included new clinical vignettes with the same 4 trajectories, introduction of an additional optional PCT test and (de)escalation opportunities. Within these new, more realistic scenarios, a negative POCT, alongside a positive PCT, did not increase antibiotic stop decisions. This study has demonstrated that contrasting test results and other behavioural factors may impair the utility of POCT. This may explain the lack of effect of POCT in improving patient outcomes (including length of hospital stay and antibiotic free days) seen in some trials investigating the use of other point-of-care biomarkers (11,22). Other studies have demonstrated the potential for PCR POCT in critical care. The INHALE WP1 study investigated two PCR-based tests for the diagnosis of pneumonia, with both tests proving to be more sensitive than routine microbiology (8). As new PCR POCTs are emerging and developing, there is a clear need for further investigation of the situations where use of POCT for infection, should be implemented (26).

Finally, the diminished effect of the negative PCR POCT in these scenarios may simply have been due to the reintroduction of further ambiguity (i.e. the borderline positive PCT). Whether ambiguity is present in the clinico-biological setting (i.e. the 2 discordant scenarios), or introduced through competing signals of the test results, it is anticipated that the requested POCT measure will act as a 'final arbiter', when

diagnostic uncertainty remains. In the higher risk setting of ICU infection decisions, most clinicians adopt caution (27). To provide clarity then, it may be sensible to choose just one rather than both point of care tests as this 'final' arbiter.

Strengths and limitations

This study offered realistic conditions and options to clinicians when making their antibiotic stop decisions. Specifically, it explored clinicians' preference for POCT vs. PCT (if any), in four different clinico-biological trajectories, whilst enabling (de)escalation options. Furthermore, we made the PCT result borderline positive, so as to 1) better represent the complexities and uncertainties of clinical practice and 2) stress-test the effect of a negative PCR POCT (i.e., assess its influence in the face of conflicting biomarker information).

We were unable to reliably analyse interim decisions, due to the small number of participants that requested either POCT or PCT. A larger sample size may have increased the size of these subgroups, allowing us to isolate and compare the respective effects of POCT and PCT on antibiotic decision making. Another way to compare their effects would be to vary their results (positive vs. negative) systematically. Presently, the effects of POCT and PCT are intertwined, which limits the conclusions that we can draw. While these conclusions are interesting and informative, further work is needed to gauge any hierarchical influence of POCT, and PCT. A direct comparison of POCT vs. PCT (with both being available but mutually exclusive) may help to answer other desired but unanswered questions; which (if any) is preferred between a direct or indirect biomarker of infection and which (if any) has a greater effect on final antibiotic decisions?

Generalizability may have been limited by the recruitment of clinicians from mainly Academic training programmes. The influence of “system-level noise” cannot also be underestimated. “Noise” refers to variation within decision making and here with regards to clinical practice (28). At a system level, variation between academic centres and hospitals (with different antibiotic prescribing policies and in-house training) can lead to noise. For example, clinicians in one hospital may be more [less] familiar with a given test than clinicians in another hospital, due to higher [lower] utilisation of that test in their hospital.

CONCLUSION

A negative POCT result does not appear to increase clinicians’ inclination to stop antibiotics, when presented alongside a borderline positive PCT test. Conflicting test results could thus be one reason why POCT has failed to increase antibiotic free days in the ICU (11,22). Further research investigating the behavioural and trajectorial factors that might compete with or override POCT in decision making, alongside initiatives to increase clinicians’ confidence in POCT, are imperative to improve its utility in the ICU. Ultimately, recognising the uncertainty in prescribing decision making, how it affects clinicians, and developing decision-making tools to support them in avoiding overreliance on antibiotics should be a future endeavour to improve antibiotic stewardship (29).

CONTRIBUTORS

SS conceived the idea. SS, MN and TL developed the protocol. MN and TL created the data collection tool. SS, MN and TL analysed the data. SS, MN and TL wrote the first and subsequent drafts. SS, MN, TL, AS, LM and NM reviewed the manuscript.

FUNDING

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

ACKNOWLEDGEMENTS

We would like to thank and acknowledge Dr Sonia Mason, Dr Marco Morosin and Dr Pascal Frederiks for their assistance in piloting the survey. In addition, we would like to express appreciation for all the clinicians, part of the WHYSTOP consortium, who participated in this study.

COMPETING INTERESTS

None declared.

DATA AVILABILITY STATEMENT

Data will be available in a public, open access repository.

REFERENCES

1. Laxminarayan R, Duse A, Wattal C, et al. Antibiotic resistance—the need for global solutions. The Lancet Infectious Diseases. 2013; 13 (12): 1057-1098. doi: 10.1016/S1473-3099(13)70318-9.
2. Vincent J, Sakr Y, Singer M, et al. Prevalence and Outcomes of Infection Among Patients in Intensive Care Units in 2017. JAMA. 2020; 323 (15): 1478-1487. doi: 10.1001/jama.2020.2717.
3. Wirz Y, Meier MA, Bouadma L, et al. Effect of procalcitonin-guided antibiotic treatment on clinical outcomes in intensive care unit patients with infection and sepsis patients: a patient-level meta-analysis of randomized trials. Critical Care (London, England). 2018; 22 (1): 191. doi: 10.1186/s13054-018-2125-7.
4. Schuetz P, Beishuizen A, Broyles M, et al. Procalcitonin (PCT)-guided antibiotic stewardship: an international experts consensus on optimized clinical use. Clinical Chemistry and Laboratory Medicine. 2019; 57 (9): 1308-1318. doi: 10.1515/cclm-2018-1181.
5. Schuetz P, Christ-Crain M, Thomann R, et al. Effect of procalcitonin-based guidelines vs standard guidelines on antibiotic use in lower respiratory tract infections: the ProHOSP randomized controlled trial. JAMA. 2009; 302 (10): 1059-1066. doi: 10.1001/jama.2009.1297.
6. de Jong E, van Oers JA, Beishuizen A, et al. Efficacy and safety of procalcitonin guidance in reducing the duration of antibiotic treatment in critically ill patients: a randomised, controlled, open-label trial. The Lancet Infectious Diseases. 2016; 16 (7): 819-827. doi: 10.1016/S1473-3099(16)00053-0.
7. Bouadma L, Luyt C, Tubach F, et al. Use of procalcitonin to reduce patients' exposure to antibiotics in intensive care units (PRORATA trial): a multicentre randomised controlled trial. The Lancet. 2010; 375 (9713): 463-474. doi: 10.1016/S0140-6736(09)61879-1.
8. Enne VI, Aydin A, Baldan R, et al. Multicentre evaluation of two multiplex PCR platforms for the rapid microbiological investigation of nosocomial pneumonia in UK ICUs: the INHALE WP1 study. Thorax. 2022; doi: 10.1136/thoraxjnl-2021-216990.

9. Poole S, Clark TW. Rapid syndromic molecular testing in pneumonia: The current landscape and future potential. The Journal of Infection. 2020; 80 (1): 1-7. doi: 10.1016/j.jinf.2019.11.021.
10. Pandolfo AM, Horne R, Jani Y, et al. Understanding decisions about antibiotic prescribing in ICU: an application of the Necessity Concerns Framework. BMJ quality & safety. 2022; 31 (3): 199-210. doi: 10.1136/bmjqs-2020-012479.
11. Hellyer TP, McAuley DF, Walsh TS, et al. Biomarker-guided antibiotic stewardship in suspected ventilator-associated pneumonia (VAPrapid2): a randomised controlled trial and process evaluation. The Lancet Respiratory Medicine. 2020; 8 (2): 182-191. doi: 10.1016/S2213-2600(19)30367-4.
12. Kostopoulou O, Mousoulis C, Delaney B. Information search and information distortion in the diagnosis of an ambiguous presentation. Judgment and Decision Making. 2009; 4 (5): 408-418. doi: 10.1017/S1930297500001236
13. Kostopoulou O, Russo JE, Keenan G, et al. Information distortion in physicians' diagnostic judgments. Medical Decision Making: An International Journal of the Society for Medical Decision Making. 2012; 32 (6): 831-839. doi: 10.1177/0272989X12447241.
14. Croskerry P. Achieving Quality in Clinical Decision Making: Cognitive Strategies and Detection of Bias. Academic Emergency Medicine. 2002; 9 (11): 1184-1204. doi: 10.1197/aemj.9.11.1184.
15. Saposnik G, Redelmeier D, Ruff CC, et al. Cognitive biases associated with medical decisions: a systematic review. BMC medical informatics and decision making. 2016; 16 (1): 138. doi: 10.1186/s12911-016-0377-1.
16. Singh S, Nurek M, Mason S, et al. WHY STOP? A prospective observational vignette-based study to determine the cognitive-behavioural effects of rapid diagnostic PCR-based point-of-care test results on antibiotic cessation in ICU infections. BMJ Open. 2023 Nov 21;13(11):e073577. doi: 10.1136/bmjopen-2023-073577.
17. Grol R, Whitfield M, De Maeseneer J, et al. Attitudes to risk taking in medical decision making among British, Dutch and Belgian general practitioners. The British Journal of General Practice: The Journal of the Royal College of General Practitioners. 1990; 40 (333): 134-136. PMID: 2115347

18. Eysenbach G. Improving the Quality of Web Surveys: The Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res 2004; 6 (3): e34. DOI: 10.2196/jmir.6.3.e34
19. Kahneman, Daniel & Tversky, Amos, 1979. " Prospect Theory: An Analysis of Decision under Risk ," Econometrica, Econometric Society, vol. 47 (2), pages 263-291, March. Handle: RePEc:ecm:emetrp:v:47:y:1979:i:2:p:263-91
20. Zeelenberg M, van den Bos K, van Dijk E, et al. The inaction effect in the psychology of regret. J Pers Soc Psychol 2002;82:314–27
21. Dhesi Z, Enne VI, O'Grady J, et al. Rapid and Point-of-Care Testing in Respiratory Tract Infections: An Antibiotic Guardian? ACS pharmacology & translational science. 2020; 3 (3): 401-417. doi: 10.1021/acsptsci.0c00027.
22. Andrews D, Chetty Y, Cooper BS, et al. Multiplex PCR point of care testing versus routine, laboratory-based testing in the treatment of adults with respiratory tract infections: a quasi-randomised study assessing impact on length of stay and antimicrobial use. BMC Infectious Diseases. 2017; 17. doi: 10.1186/s12879-017-2784-z.
23. Singhal A, Nurek M, Lau T et al. Infection Point of Care Tests (POCT) in simulated vignettes can improve the education of Antibiotic Stewardship Programmes. An observational comparative study of Clinicians vs Medical Students. Research Square [Preprint] September 22, 2023. https://doi.org/10.21203/rs.3.rs-3295414/v1
24. Trimble M, Hamilton P. The thinking doctor: clinical decision making in contemporary medicine. Clinical Medicine. 2016; 16 (4): 343-346. doi: 10.7861/clinmedicine.16-4-343.
25. Croskerry P. Clinical cognition and diagnostic error: applications of a dual process model of reasoning. Advances in Health Sciences Education: Theory and Practice. 2009; 14 Suppl 1 27-35. doi: 10.1007/s10459-009-9182-2.
26. Newcombe V, Coats T, Dark P, et al. The future of acute and emergency care. Future Healthcare Journal. 2021; 8 (2): e230-e236. doi: 10.7861/fhj.2021-0097.
27. Pandolfo AM , Horne R , Jani Y, et al. Understanding decisions about antibiotic prescribing in ICU: an application of the necessity concerns framework. BMJ Qual Saf 2022;31:199–210. doi:10.1136/bmjqs-2020-012479

28. Kahneman D, Sibony O, Sunstein C. Noise: A Flaw in Human Judgement. London: William Collins; 2021.
29. Tarrant C, Krockow EM. Antibiotic overuse: managing uncertainty and mitigating against overtreatment. BMJ Qual Saf 2022;31:163–167. doi:10.1136/bmjqs-2021-013615

FIGURE LEGENDS

Figure 1. Number and proportion of decisions made initially (before receiving any test results) vs. finally (after receiving the negative POCT and positive PCT results), regardless of scenario (n=264). Participants chose between the options of escalate (more than the original course), continue with the original course, de-escalate (less than the original course) and stop with regards to antibiotics.

Figure 2. Number of clinicians that chose to STOP antibiotics initially (i.e., before receiving any test results) vs. finally (after receiving both a negative POCT result and a positive PCT result), per scenario. The total number of responses in each scenario (improvement, disc clin better, disc clin worse, and worsening) was 66. * $p<0.05$

Figure 3. Median willingness-to-stop before vs. after receiving the negative POCT and positive PCT results, per scenario (n=66 for each scenario). Willingness-to-stop represents a participant's confidence (1-6), signed positive if they chose to stop antibiotics and negative if they chose to continue (be it via continuation of the original course, escalation, and de-escalation). * $p<0.05$

Figure 4. Number of participants requesting POCT and/or PCT tests, per scenario (n=66 for each scenario). Participants were offered a POCT and PCT test within each scenario, after making an initial antibiotics decision, and could request POCT only, PCT only, both tests or neither test.

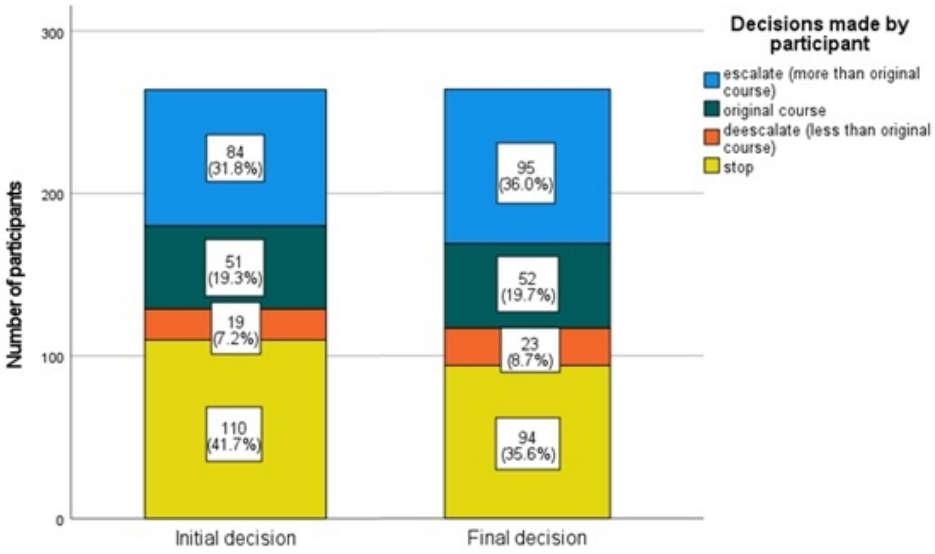


Figure 1. Number and proportion of decisions made initially (before receiving any test results) vs. finally (after receiving the negative POCT and positive PCT results), regardless of scenario (n=264). Participants chose between the options of escalate (more than the original course), continue with the original course, de-escalate (less than the original course) and stop with regards to antibiotics.

155x96mm (96 x 96 DPI)

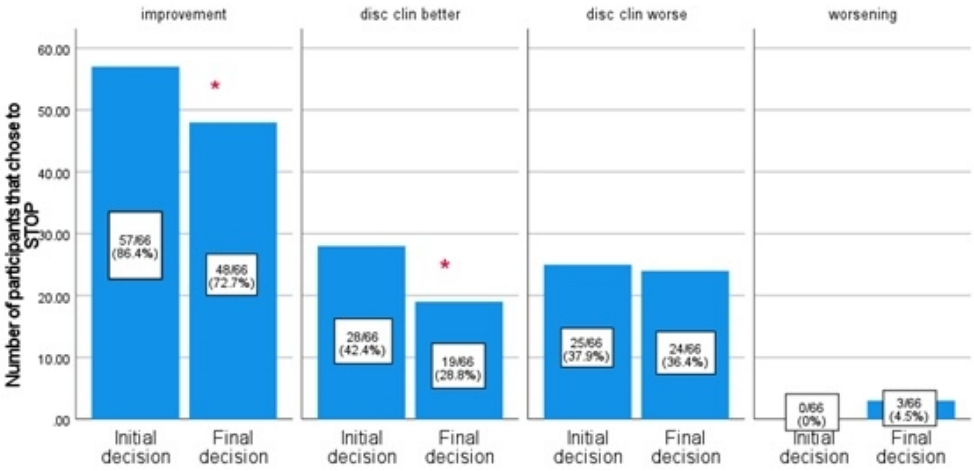


Figure 2. Number of clinicians that chose to STOP antibiotics initially (i.e., before receiving any test results) vs. finally (after receiving both a negative POCT result and a positive PCT result), per scenario. The total number of responses in each scenario (improvement, disc clin better, disc clin worse, and worsening) was 66. * p<0.05

150x88mm (96 x 96 DPI)

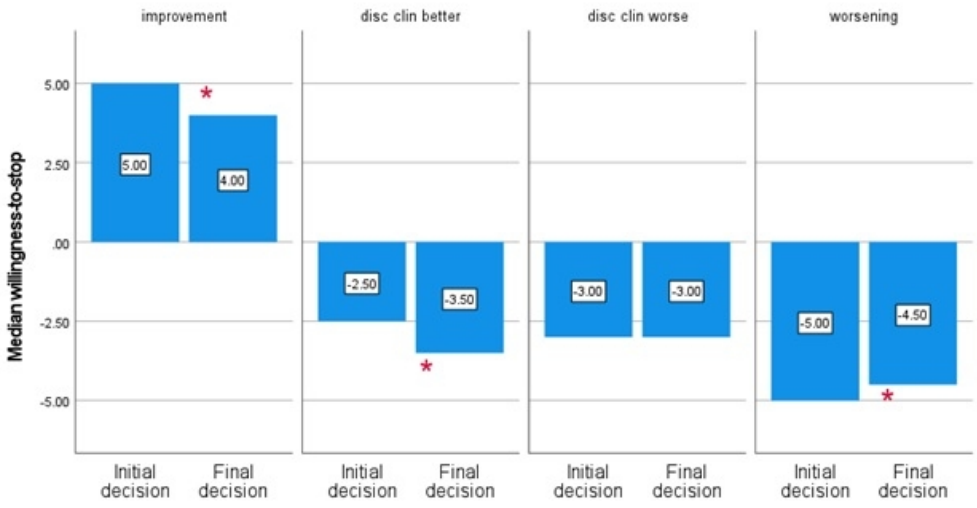


Figure 3. Median willingness-to-stop before vs. after receiving the negative POCT and positive PCT results, per scenario (n=66 for each scenario). Willingness-to-stop represents a participant's confidence (1-6), signed positive if they chose to stop antibiotics and negative if they chose to continue (be it via continuation of the original course, escalation, and de-escalation). * $p<0.05$

165x97mm (96 x 96 DPI)

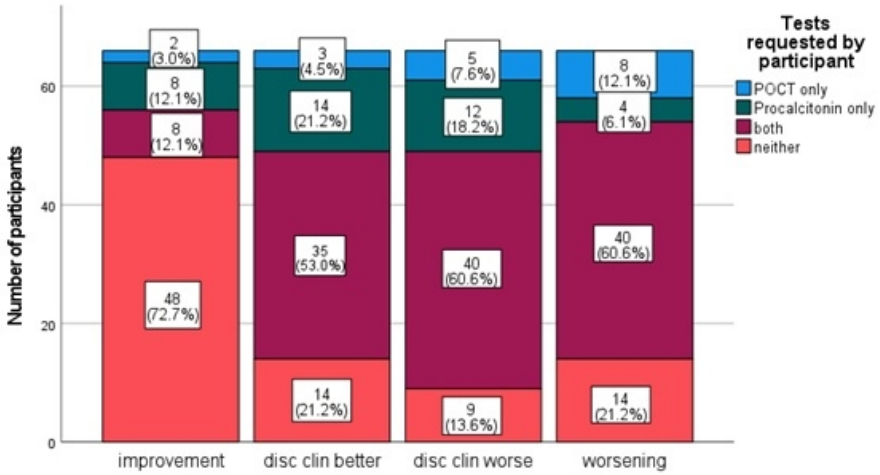


Figure 4. Number of participants requesting POCT and/or PCT tests, per scenario (n=66 for each scenario). Participants were offered a POCT and PCT test within each scenario, after making an initial antibiotics decision, and could request POCT only, PCT only, both tests or neither test.

165x97mm (96 x 96 DPI)

SUPPLEMENTAL MATERIAL

SM1

Sample size calculation.

The original WHYSTOP study found that a negative POCT result increased clinicians' willingness to stop antibiotics significantly ($p<0.01$). Specifically: prior to receiving the negative POCT result, clinicians were willing to stop antibiotics 54% of the time (138/258); after receiving the result, they were willing to stop antibiotics 70% of the time (180/258; chi-square=25.82, df=1, $p<0.01$, $w=0.32$). Using G*Power 3.1, we estimated that a minimum of 77 responses would be required to replicate this effect, with power at 80%, alpha at 0.05, and 1 degree of freedom. To account for clustered data (with each participant seeing 4 scenarios), we then calculated the "design effect" (DE)(1), using the formula $1+(n-1)\rho$, where n is the cluster size (4) and ρ the intraclass correlation/Cronbach's alpha (2) from Singh et al.'s study (0.061). Multiplying the number of required responses (77) by the DE (1.183) suggested that 91 responses were needed (77 x 1.183). At 4 responses per participant, 23 participants were required (91/4).

1. Donner A, Klar N. *Design and Analysis of Cluster Randomization Trials in Health Research.* London: Arnold; 2000. Available from: https://www.jameslindlibrary.org/donner-a-klar-n-2000/ [Accessed May 24, 2022].
2. Bi J, Kuesten C. Intraclass Correlation Coefficient (ICC): A Framework for Monitoring and Assessing Performance of Trained Sensory Panels and Panelists. *Journal of Sensory Studies.* 2012; 27 (5): 352-364. doi: 10.1111/j.1745-459X.2012.00399.x.

SM2a

The *improvement* vignette

Details of admission:

A 68-year-old male presented following a fall at home. He sustained rib fractures to his right anterior 5th and 6th ribs and was admitted for pain control. He has a background of poorly controlled insulin-dependent type 2 diabetes. Two days into admission, he developed hypoxia and pyrexia. His observations were the following:

Respiratory rate	26/min
SpO2	90% on room air
Heart rate	90/min sinus rhythm
Blood pressure	111/81 mmHg
Temperature	38.0 C

There was right sided consolidation on his chest radiograph and nil else. His blood tests demonstrated a WBC of 15 and CRP of 78. He was initiated on Co-amoxiclav.

Five days later:

7 days into his admission (5 days following antibiotics), he had improved shortness of breath and was afebrile. His observations were:

Respiratory rate	18/min
SpO2	99% on room air
Heart rate	83/min sinus rhythm
Blood pressure	112/80 mmHg
Temperature	37.0 C

He was pain free and mobilising on the ward. His repeat blood tests demonstrated a WBC of 8 and CRP of 15.

SM2b

The *overall worse* vignette

Details of admission:

A 78-year-old male was admitted with a 4-day history of worsening shortness of breath and a productive cough. He has a background of hypertension, Type II Diabetes Mellitus and a previous TIA (2019). He has no known drug allergies. His admission observations were:

Respiratory rate	22/min
SpO2	87% on room air
Heart rate	101/min sinus rhythm
Blood pressure	106/62 mmHg
Temperature	37.9 C

There was right basal consolidation on his chest radiograph. His blood tests demonstrated a WBC of 12 and a CRP of 70. He was empirically started on Levofloxacin and Clarithromycin. Within 24 hours he deteriorated and required mechanical ventilation.

Five days later:

5 days into his admission and after an initial improvement in ventilation, he became febrile. His observations were:

Respiratory rate	25/min
SpO2	90% on FiO2 21%
Heart rate	120/min sinus rhythm
Blood pressure	110/58 mmHg
Temperature	38.9 C

Further clinical assessment does not identify an alternative source of infection. However, his repeat blood tests demonstrated a WBC of 15 and a CRP of 150.

SM2c

The *disc clin better* vignette

Details of admission:

A 60-year-old male was admitted to ITU with a 6-day history of pyrexia, shortness of breath, and a productive cough with rusty sputum. He has a past medical history of well controlled, uncomplicated HIV (CD4 count >500 and viral load undetectable 3 months ago). His observations were the following:

Respiratory rate	32/min
SpO2	80% on room air
Heart rate	115/min sinus rhythm
Blood pressure	95/52 mmHg
Temperature	38.4 C

A chest radiograph demonstrated left midzone and basal consolidation. His blood tests demonstrated a WBC of 13 and a CRP of 50. Sputum culture grew MRSA. He was intubated and ventilated and empirically started on Linezolid. His initial blood gas findings were:

- FiO2 0.6
- PaO2 of 7.7 kPa
- PaCO2 5.2 kPa
- Base excess of -4

Seven days later:

7 days into his admission, he was improving on ventilation and extubated and weaned onto room air. He was feeling notably better. Clinical assessment does not identify any clinical source of infection. His observations were:

Respiratory rate	18/min
SpO2	96% on room air
Heart rate	82/min sinus rhythm
Blood pressure	124/70 mmHg
Temperature	36.9 C

However, his repeat blood tests demonstrated a worsening with a WBC of 15 and a CRP of 60.

SM2d

The *disc clin worse* vignette

Details of admission:

A 59-year-old male was admitted to ITU with a 3-day history of vomiting, pyrexia and a productive cough. He has a background of alcohol excess. His observations were the following:

Respiratory rate	30/min
SpO2	88% on room air
Heart rate	130/min sinus rhythm
Blood pressure	115/64 mmHg
Temperature	38.3 C

A chest radiograph demonstrated bilateral pulmonary infiltrates. His blood tests demonstrated a WBC of 20 and a CRP of 82. He was intubated and ventilated and empirically started on Piperacillin/Tazobactam. His initial blood gas findings on ventilation were:

- FiO2 0.5
- PaO2 of 8.2 kPa
- PaCO2 4.5 kPa
- Base excess of -5

Five days later:

5 days into his admission, he was making a good recovery, onto low pressure support ventilation and FiO2 down to 0.3. Chest radiograph findings were unchanged at this point. However, 1 day later, he developed new pyrexia to 37.8 C and increased oxygen requirement to FiO2 0.45. Investigations ruled out a pulmonary embolus.

Seven days later:

7 days into his admission, his observations were:

Respiratory rate	20/min
SpO2	92% on FiO2 0.45
Heart rate	100/min sinus rhythm
Blood pressure	130/70 mmHg
Temperature	37.8 C

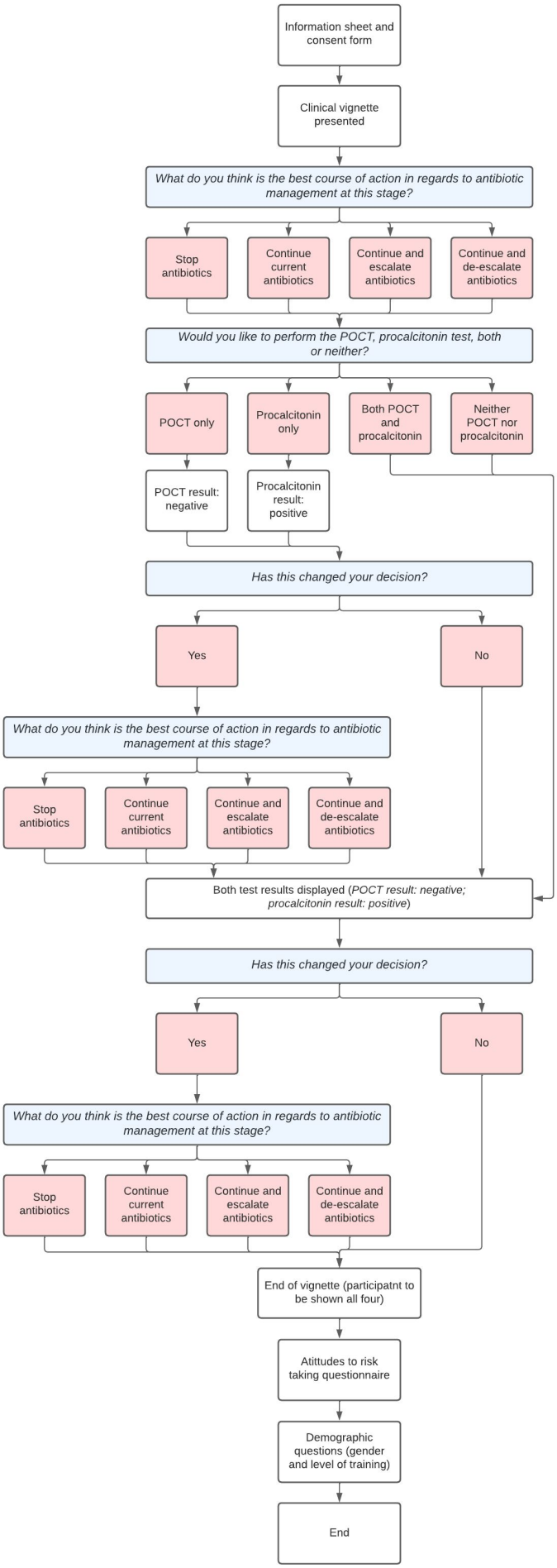
Clinical assessment does not identify an alternative source of infection.

His repeat blood tests demonstrated an ongoing reduction with a WBC of 10 and a CRP of 12.

SM3. Reasons for Clinicians' choice of diagnostic test when offered.

Reasons for performing POCT only	Reasons for performing PCT only	Reasons for performing both POCT and PCT	Reasons for performing neither POCT nor PCT

• I trust the POCT; • The POCT is necessary in this case; • I feel confident interpreting the POCT results; • I do not trust the procalcitonin test (I have concerns regarding the accuracy of the test); • The procalcitonin test is unnecessary in this case; • I do not feel confident interpreting the procalcitonin test results; • Other (if selected, the participant was asked to elaborate using free text).	• I trust the PCT test; • The PCT test is necessary in this case; • I feel confident interpreting the procalcitonin test results; • I do not trust the POCT (I have concerns regarding the accuracy of the test); • The POCT is unnecessary in this case; • I do not feel confident interpreting the POCT; • Other (if selected, the participant was asked to elaborate using free text).	• To supplement my clinical judgement; • I trust these tests; • The tests are necessary in this case; • I feel confident interpreting these tests; • Other (if selected, the participant was asked to elaborate using free text).	• I prefer to rely on my clinical judgement; • I do not trust these tests; • These tests are unnecessary in this case; • I don't feel confident interpreting these tests; • Other (if selected, the participant was asked to elaborate using free text).



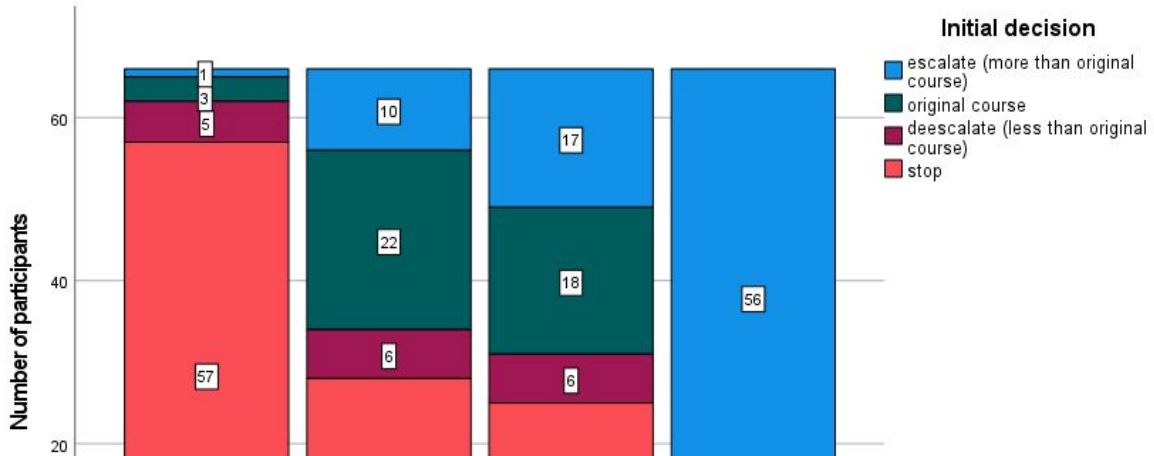
SM4. Graphical representation of the survey flow and procedure. Blue boxes indicate key questions, with the boxes below (red) displaying the possible responses. Created using Lucidchart (Lucid Software Inc., Utah, USA).

SM5. Piloting process

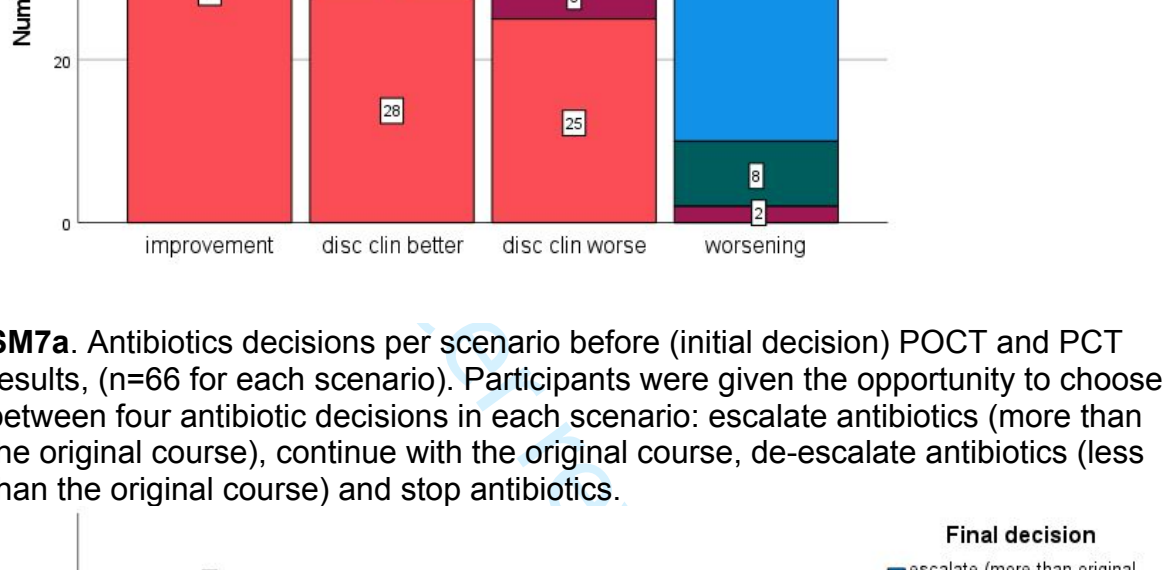
The survey was constructed and tested between the authors before piloting began. Two non-participating intensive care clinicians (SpR trainees) known to the authors were recruited to pilot-test the vignettes and survey. Feedback was given regarding the clarity and accessibility of the survey, as well as its format and structure. Feedback was very positive regarding the structure of the survey and contents of the vignettes, with only minor changes made to the survey. Particularly, we inserted a statement within the vignettes to suggest that there was no sign of alternative infection, as was suggested. Following this, the survey was trialed on other non participating ICU clinicians, then finalized and participant recruitment began.

SM6. Demographic and experience characteristics of the sample (n = 66).

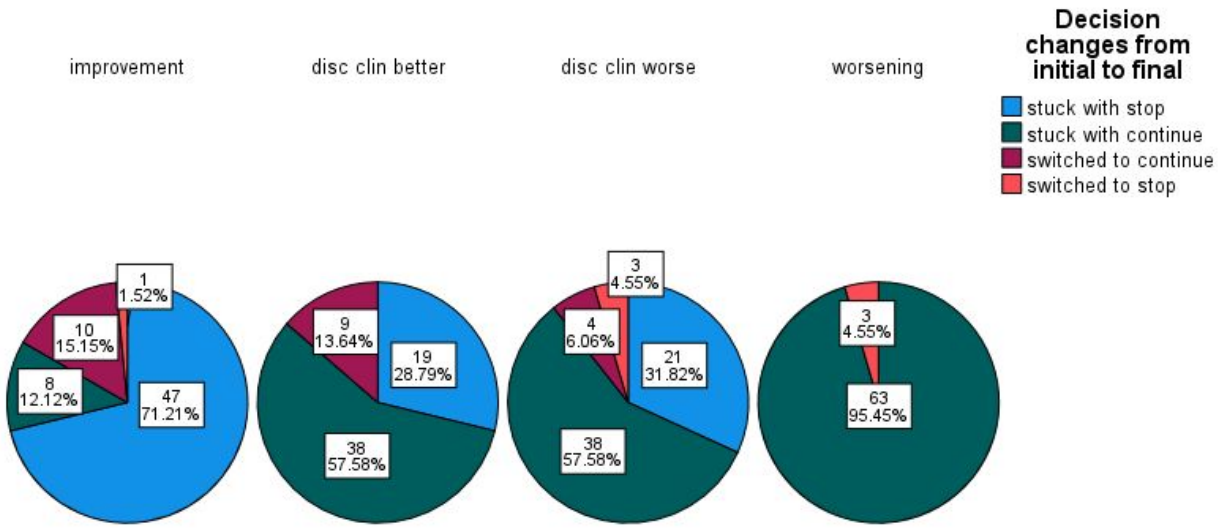
	n (%)	***Mean (SD), range***
Gender		
Male	39 (59.1%)	
Female	25 (37.9%)	
Prefer not to say	2 (3.0%)	
Grade		
Consultant	34 (51.5%)	
SpR trainee	17 (25.8%)	
SHO trainee	13 (19.7%)	
FY trainee	2 (3.0%)	
Experience: consultants		
Number of years since consultancy awarded		9.77 (7.67), 0 - 25
Experience: trainees		
>24 months on ICU ward	12 (37.5%)	
12-24 months on ICU ward	5 (15.6%)	
6-12 months on ICU ward	3 (9.4%)	
3-6 months on ICU ward	12 (37.5%)	



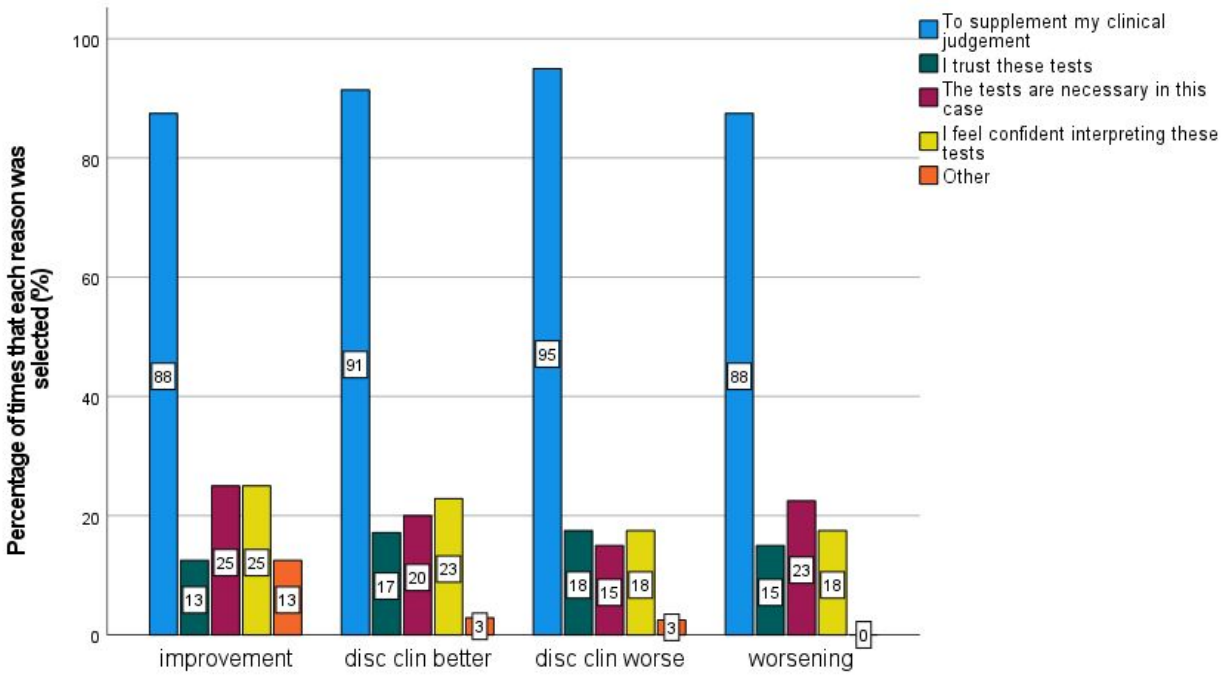
SM7a. Antibiotics decisions per scenario before (initial decision) POCT and PCT results, (n=66 for each scenario). Participants were given the opportunity to choose between four antibiotic decisions in each scenario: escalate antibiotics (more than the original course), continue with the original course, de-escalate antibiotics (less than the original course) and stop antibiotics.



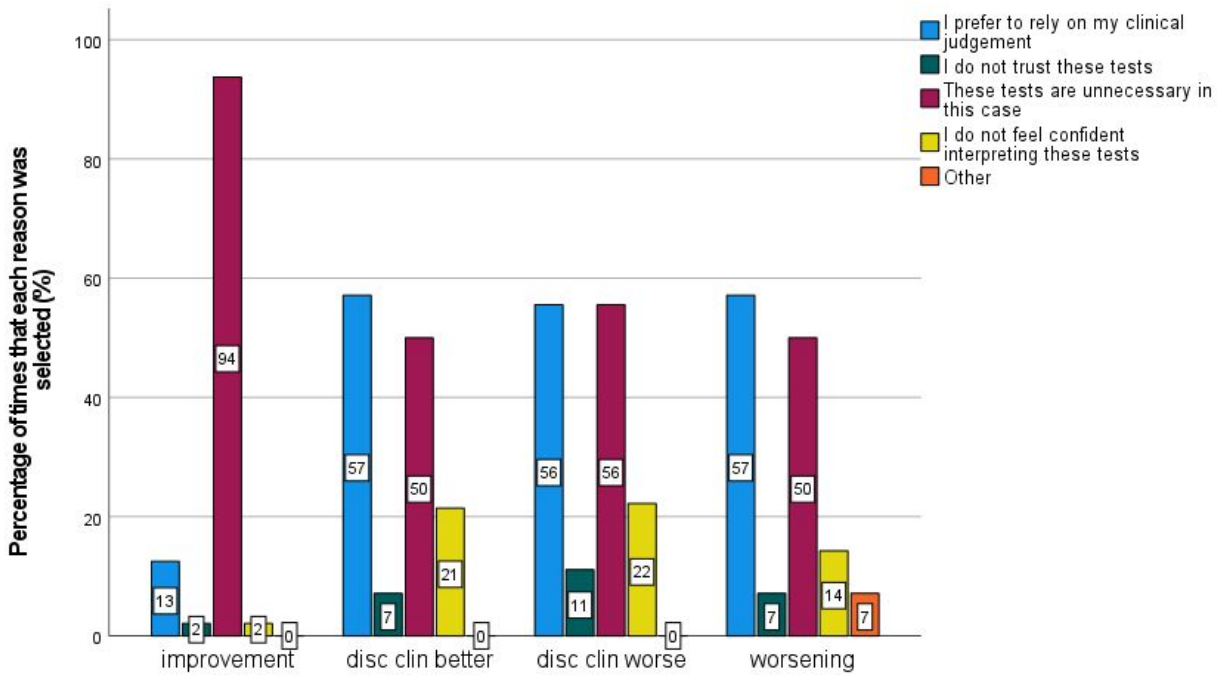
SM7b. Antibiotics decisions per scenario after POCT and PCT results (final decision), (n=66 for each scenario). Participants were given the opportunity to choose between four antibiotic decisions in each scenario: escalate antibiotics (more than the original course), continue with the original course, de-escalate antibiotics (less than the original course) and stop antibiotics.



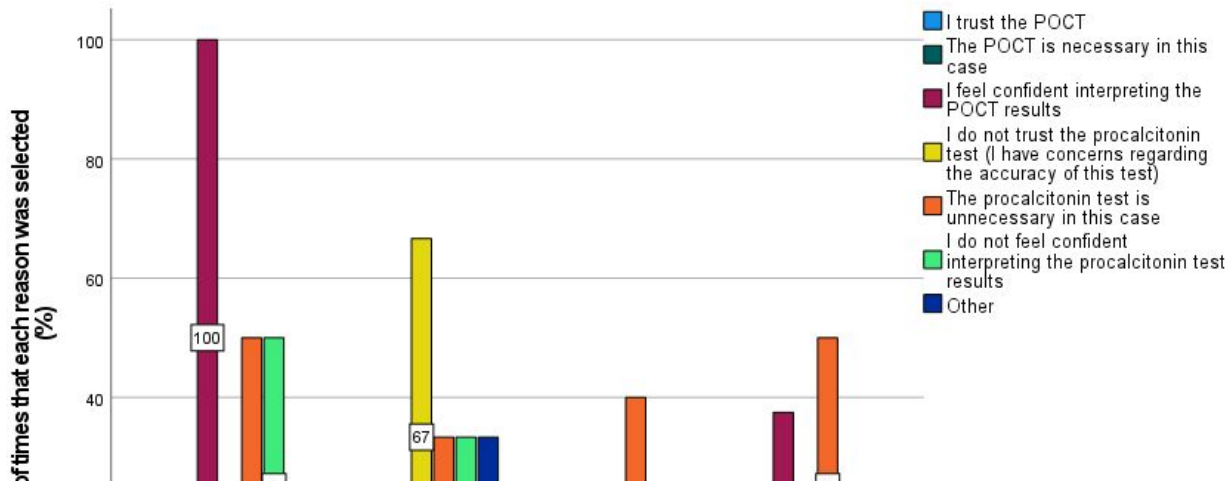
SM8. Decision changes, before vs. after receiving the negative POCT and positive procalcitonin test results, per scenario (n=66 for each scenario). Decisions were classified as “continue” if the participant elected to continue with the original course, escalate, or de-escalate. were all classified as “continue”. Absolute numbers and percentages are shown for each scenario.



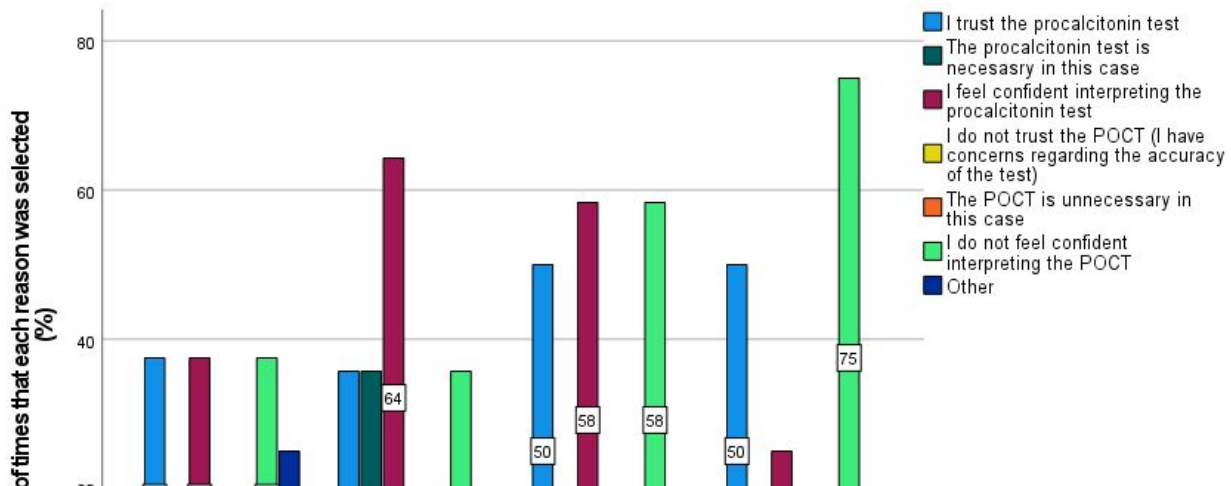
SM9a. Reasons for decision to request **neither** POCT nor PCT given by clinicians in the improvement (n=48), disc clin better (n=14), disc clin worse (n=9) and worsening (n=14) scenarios. Participants were able to select more than one reason for their decision. Data labels indicate the percentage that each reason was selected within each scenario.



SM9b. Reasons for decision to request **both** the POCT and PCT given by clinicians in the improvement (n=8), disc clin better (n=35), disc clin worse (n=40) and worsening (n=40) scenarios. Participants were able to select more than one reason for their decision. Data labels indicate the percentage that each reason was selected within each scenario.



SM9c. Reasons for decision to request the **POCT only** given by clinicians in the improvement (n=2), disc clin better (n=3), disc clin worse (n=5) and worsening (n=8) scenarios. Participants were able to select more than one reason for their decision. Data labels indicate the percentage that each reason was selected within each scenario.



SM9d. Reasons for decision to request the **PCT only** given by clinicians in the improvement (n=8), disc clin better (n=14), disc clin worse (n=12) and worsening (n=4) scenarios. Participants were able to select more than one reason for their decision. Data labels indicate the percentage that each reason was selected within each scenario.

SM10. A mixed effects ordinal logistic regression model of final antibiotics decision on 1) initial antibiotics decisions, 2) test(s) requested, 3) attitudes towards risk taking, 4) level of experience (consultant vs. trainee), and 5) scenarios returned the following results. A p-value of less than 0.05 was considered as statistically significant.

Parameter	***OR***	**95% Confidence Interval**	***p*-value**
Initial antibiotics decisions	8.70	[5.40 to 14.02]	<0.001
POCT request	3.84	[1.61 to 9.15]	0.002
PCT request	0.26	[0.12 to 0.55]	<0.001
Experience	1.56	[0.85 to 2.86]	0.156
Sum of attitude toward risk taking score	1.02	[0.92 to 1.12]	0.761
Scenario	0.63	[0.32 to 1.22]	0.168

SM11. A mixed effects linear regression model of final willingness-to-stop on 1) initial willingness-to-stop, 2) test(s) requested, 3) attitudes towards risk taking, 4) level of experience (consultant vs. trainee), and 5) scenarios returned the following results. A p-value of less than 0.05 was considered as statistically significant.

Parameter	***b***	**95% Confidence Interval**	***p*-value**
Initial willingness-to-stop	0.70	[0.57 to 0.83]	<0.001
POCT request	1.41	[0.15 to 2.66]	0.028
PCT request	-1.54	[-2.71 to -0.38]	0.009
Experience	0.37	[-0.66 to 1.41]	0.483
Sum of attitude toward risk taking score	0.07	[-0.10 to 0.24]	0.418
Scenario	-0.32	[-1.10 to 0.46]	0.418
Constant	-0.90	[-3.41 to 1.61]	0.484

Item category	Checklist item	Page no.	Description
Design	Study design	2, 7	Prospective observational study. Simulated clinical scenarios.
	Ethics approval	13	Institutional Review Board approval.
	Informed consent	10	As part of the online Questionnaire. The survey began with an information sheet and consent form, which explained (inter alia) the length of time of the survey, which data were stored and where and for how long, who the investigators were and the purpose of the study. After the reading these, participants to provide informed consent by ticking a box labelled "I agree to participate in this study". If they did not tick this box, they were unable to proceed with the survey.
	Data protection		Personal information was not collected or stored as part of the survey. Participants were invited (but not required) to join our consortium ("WHY STOP"); those that elected to join were directed to a separate survey, where they were asked to provide their name and email address. These data were stored securely on a password-protected university computer. Survey responses remained wholly anonymous.
Development and pre-testing		11	This is explained in the manuscript methods and supplementary materials SM5. In an iterative process, surveys were repeatedly piloted and revised to ensure usability, technical functionality and plausibility of the clinical scenarios used.
Recruitment process	Open v Closed survey	7-8	The survey was accessible to consultants and trainees in Intensive Care Medicine (ICM; 3+ months continuous experience in ICU) currently working in London-based university teaching hospitals

	Contact mode	7-8	Initial contact was made via internet portals/forums and WhatsApp groups. Clinicians who took part in a previous survey and had indicated a willingness to be involved in future studies were contacted via direct email.
	Advertising the survey	7-8	The study was advertised via known trainee and consultant training portals (local), social media, and word of mouth.
Survey administration	Web/email	8	The survey was hosted on the secure Qualtrics platform and accessible via a website. Responses were downloaded from Qualtrics (i.e., captured automatically, not entered manually).
	Context	8	The website is the homepage of the WHY STOP Consortium, a diverse group of clinicians and behavioural scientists that aim to improve antibiotic stewardship in ICU (why-stop.wixsite.com/itu-decision-making). The website contains information about the Consortium, the core research team, and any past/present studies. For those who wish to participate in an open study, the website provides information on how to do so (e.g., a link to an open questionnaire). There was nothing on the website that might influence the selection of participants.
	Mandatory/voluntary	7-8	This was a voluntary study.
	Incentives	8	Study participants were invited to contribute to the manuscript, and to have this contribution acknowledged within the manuscript. This was entirely optional.
	Time/date	8	March-Oct 2022
	Item randomisation	10	Clinical vignettes were presented in a random order.
	Adaptive questioning	10-11	Clinicians were asked what tests they would like to request and could choose between POCT only, PCT only, both POCT and PCT, neither POCT nor PCT. They were then asked to

			indicate their reason/s by selecting from a list. These lists were different depending on the answer selected.
	Number of items	10-11	13 per vignette. At the end of the survey, participant's completed a measure of individual differences (5 items) and gave demographic information (5 items)
	Number of screens		25 in total (5 per patient case plus 5 for information sheet, consent form, introduction, risk inclination questionnaire and demographic questions).
	Completeness check		Mandatory items were enforced via Qualtrics' inbuilt functionality (participants could not proceed if they did not respond to mandatory items). Participants could select "prefer not to say" for demographic questions such as gender. Completeness was checked during piloting with junior doctors who were not involved in the actual survey.
	Review step		Participants were given the opportunity to update their answers during the course of the patient scenarios. This was an integral part of the study design.
Response rates	Unique site visitor		Each participant had a unique identifier as determined by the Qualtrics system ("ResponseID"). This was based on cookies (not IP Addresses).
	View rate		75 doctors started the survey. Of these, 66 completed all 4 scenarios
	Participation rate		75 doctors started the survey. Of these, 66 completed all 4 scenarios.
	Completion rate		75 doctors started the survey. Of these, 66 completed all 4 scenarios
Preventing multiple entries for same individual	Cookies used		When respondents started a session, Qualtrics placed a cookie on their browsers that kept track of their survey progress. Respondents had one week to return to the survey and finish their response. After a week, the response was recorded "as is". We made no attempt to prevent

			duplicate entries, given that clinicians had zero incentive (financial or otherwise) to do this and survey completion time was lengthy (20-25min).
	IP check		We made use of Qualtrics' "Anonymize responses" option, whereby client IP Addresses are not collected or stored. We did not check for duplicate responses, for the reasons stated above.
	Log file analysis		We did not check for duplicate responses, for the reasons stated above.
	Registration		The survey was accessed through an open link. We cannot be 100% sure that participants didn't fill in the form in duplicate. However, we think this highly unlikely given the length of the survey, the incentive structure (which did not change with submission of multiple surveys), and the degree of trust we maintain in colleagues.
Analysis	Handling of incomplete questionnaires	13	Incomplete responses were not utilised or analysed as they would be automatically removed after 1 week. Therefore, if a participant completed 3 scenarios and stopped mid-way through the 4th, these scenarios would not be used.
	Questionnaires with atypical timestamp		The minimum time needed to read and respond to a single scenario was 90 seconds (measured during piloting). We aimed to exclude respondents that took less time than this on a given scenario, but none did.
	Statistical correction	11-12	Not applicable. All statistical operations are described in the Analyses section of Methods.

BMJ Open

Point of Care Tests, Diagnostic Uncertainty and Antimicrobial Stewardship in the ICU. Procalcitonin or PCR to aid antibiotic stop decisions; an observational cohort study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2024-084872.R1
Article Type:	Original research
Date Submitted by the Author:	18-Oct-2024
Complete List of Authors:	Lau, Timothy; Imperial College London, Nurek, Martine; Imperial College London, Surgery and Cancer Singhal, Archit; Imperial College London Moore, Luke; Chelsea and Westminster Hospital NHS Foundation Trust, Department of Microbiology; Imperial College London, National Institute for Health Research Health Protection Research Unit in Healthcare Associated Infections and Antimicrobial Resistance Mughal, Nabeela; Imperial College London; Chelsea and Westminster Hospital NHS Foundation Trust Singh, Suveer; Imperial College London, Faculty of Medicine; Chelsea and Westminster Hospital NHS Foundation Trust, Respiratory and Intensive Care Medicine
Primary Subject Heading:	Intensive care
Secondary Subject Heading:	Infectious diseases
Keywords:	Behavior, MICROBIOLOGY, INFECTIOUS DISEASES, INTENSIVE & CRITICAL CARE, Diagnostic microbiology < INFECTIOUS DISEASES

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Point of Care Tests, Diagnostic Uncertainty and Antimicrobial Stewardship in the ICU. Procalcitonin or PCR to aid antibiotic stop decisions; an observational cohort study

Lau, Timothy[1,§], Nurek, Martine[1,§], Singhal, Archit[1], Moore, Luke[1,2], Mughal, Nabeela[1,2], Singh, Suveer[1,2,§,*]

[1] Imperial College London, London, UK

[2] Chelsea and Westminster Hospital NHS Foundation Trust, London, UK

[§] These authors contributed equally to this work

*Correspondence to:

Professor Suveer Singh

Departments of Adult Intensive Care Medicine and Respiratory Medicine, APMIC, Imperial College London, 369 Fulham Road, London SW10 9NH, UK

suveer.singh@imperial.ac.uk

WORD COUNT: 4855

ABSTRACT

OBJECTIVES: ICU clinicians stop antibiotics more often, with a negative infection-point of care test (PCR-POCT). Simulated cases of diagnostic uncertainty regarding infection resolution, led clinicians to choose options such as Procalcitonin (PCT) and/or PCR-POCT tests +\-de/escalation to aid stop decisions. We hypothesized that a direct infection indicator, PCR-POCT would influence stop judgements moreso than indirect PCT. Accordingly, we tested antibiotic stop decisions when presented a negative PCR-POCT, despite borderline positive PCT.

DESIGNS: Observational prospective study

SETTING: Intensive Care Unit

PARTICIPANTS: 66 ICU clinicians from University hospitals.

METHODS: Clinicians saw 4 scenarios of different clinico-biological trajectories:1) clear improvement, 2) clear worsening, 3) discordant-clinically better/biologically worse, 4) discordant-clinically worse/biologically better. Participants gave an initial decision (stop/ continue/continue-escalate/continue-de-escalate). Then PCR-POCT and/or PCT was offered (accept/decline). Irrespective, after a negative PCR-POCT and borderline positive PCT result, a final antibiotic decision was made.

MEASURES: Proportion of stop decisions before vs. after test results per scenario. The association of the final decision with the clinician's change in confidence, willingness to request the biomarker(s), and the case trajectory were determined.

RESULTS: Fewer clinicians than expected stopped antibiotics v baseline (36%, 94/264 vs 42%, 110/264, p=0.045). This was so in 3 of 4 scenarios, significantly less in the improvement (p<0.001) and the discordant clin-better scenario (p=0.024). PCT was requested more frequently than PCR-POCT (61% vs. 53%, p<0.001). PCT requesters (v declining) were significantly less inclined to stop antibiotics (p<0.001),

whilst PCR-POCT requesting led to more stopping (p<0.001), before knowing test results.

CONCLUSIONS: A negative PCR-POCT result did not increase clinicians' inclination to stop antibiotics, when alongside a borderline positive PCT. This reflects clinicians' natural risk aversion. PCT was more popular than PCR-POCT, but PCR-POCT was more likely to aid stop decisions.

Their comparison, role, utility and selective deployment for influencing antibiotic stop decisions more effectively requires a large RCT.

STRENGTHS AND LIMITATIONS OF THIS STUDY

- In four typical clinical vignettes, with different clinico-biological trajectories, this study offered a realistic simulation of ICU-related respiratory infection factors to test antibiotic stewardship decisions.
- Choices to escalate/de-escalate antibiotics alongside stop and continue options, provide a reproducible and adaptable test platform to study the clinical situational and behavioural factors that influence antibiotic decisions.
- By focussing on the end (rather than onset) stage of infection, it offered the opportunity to test clinicians' preferences for direct (PCR-POCT) or indirect (PCT) point of care tests as arbiters for antibiotic decisions, when confronted with clinic-biological diagnostic uncertainty.
- A larger sample is required to robustly determine preferences between PCR-POCT or PCT, and influential factors. This would have to utilize the same choices (e.g. continue/stop/(de)escalate), whilst adding the presentation of theoretical combinations of positive and negative results, using the same test vignette platform.

BACKGROUND

Antimicrobial resistance has become an increasingly pertinent issue within patient care [1]. Antibiotic stewardship programmes (ASPs) are strategies that aim to improve the use of antibiotics and have been employed to reduce the potential for antimicrobial resistance. Antibiotic prescribing is high in the ICU setting, with up to 70% of patients having an antibiotic prescribed and so the use of ASPs in this setting can have large effects on antimicrobial resistance [2]. Some ASPs have had success in reducing antibiotic usage, including procalcitonin (PCT)-guided antibiotic stewardship for sepsis in the ICU [3]. PCT is a surrogate biomarker for infection and is considered a useful tool in antibiotic prescribing, having been successful in reducing antibiotic course lengths in study settings [4-7]. Another useful tool to potentially improve antibiotic prescribing is the polymerase chain reaction point of care test (PCR-POCT, herein also referred to as POCT). These tests have been used during the COVID-19 pandemic and are found to have high diagnostic accuracy [8]. Infection identifying POCTs can effectively rule out the presence of infective organisms; thereby increasing clinicians' confidence to stop antibiotics [9]. This is important as clinicians are found to continue (rather than stop) empirical antibiotics when there is clinical uncertainty, based on natural risk averseness, particularly with more severe illness [10]. POCT can alleviate this uncertainty and should therefore reduce antibiotic prescribing. However recent studies have not seen this. In the VAPrapid trial of ventilator associated pneumonia, use of a highly accurate cytokine based POCT (interleukin 1/8) rule out test failed to increase antibiotic-free days in ICU despite excellent test performance [11].

One possible explanation is that factors such as cognitive biases, may be impacting the use of these tests and antibiotic decision making more broadly. Confirmation

bias, whereby clinicians find and interpret evidence to support an existing hypothesis, could affect decision making [12,13]. Anchoring, whereby clinicians fixate on early information in a case, may lead clinicians to overlook important possibilities [14]. While these clinician factors have been acknowledged in previous work, research investigating their specific and quantifiable effect on decision making is lacking [15].

A recent study investigated clinical and clinician factors that influence POCT-use and antibiotic prescribing in ICU [16]. This vignette-based experiment found that a negative PCR-POCT result (suggesting no infection) significantly increased clinicians' inclination to stop antibiotics, but three "competing" factors worked to decrease it: an ambiguous or worsening patient trajectory, clinicians' first impressions (i.e., high confidence that antibiotics were needed) and lack of interest in POCT (i.e., rejection of the POCT when it was offered). Whilst that study highlighted the potential utility of POCT for antibiotic cessation in the setting of ICU related respiratory infection, as well as the factors that might diminish its utility, the model was simple with limitations. Four vignettes describing patients that had just completed a course of antibiotics were constructed: one was clearly improving (clinical and biological signs better than upon admission), one was clearly deteriorating (clinical and biological signs worse), and two were ambiguous (clinical signs better but biological signs worse / clinical signs worse but biological signs better). The study found that POCT was most requested and most effective (in promoting antibiotic stop decisions) in the ambiguous scenarios that featured clinical worsening but biological improvement (and overall worsening). This is important, as a critical component of increasing and streamlining POCT use in the ICU is identifying scenarios in which it is most/least helpful. For reliability, those findings

would benefit from replication in a new set of vignettes that adhere to the same structure (i.e., a new vignette depicting clear improvement, a new vignette depicting clear worsening, etc).

In the original study, participants' uptake of the POCT offer was relatively high (overall 65% of the time). In clinical practice, however, clinicians have other tests available, which are more established and therefore more widely used than POCT (e.g., procalcitonin, PCT). Whether clinicians would still request POCT when PCT is available (albeit an indirect biomarker surrogate of an infectious agent rather than a direct form in PCR-POCT) remains to be seen.

Thirdly, in the original study, clinicians were only asked to choose between two courses of action: stop or continue antibiotic treatment, in reality, the opportunity to (de)escalate, are further options. Recognising this limitation, the authors asked participants at the end of the study: "*Had the option to de/escalate antibiotics been available, would you have used it?*" Most clinicians (74.3%, 52/70) responded yes, leading the authors to conclude that findings may have looked different had (de)escalation been available.

The present study aimed to address this limitation. Specifically, to replicate and validate the previous findings, by making three improvements to the study design:

1. clinicians were presented with a new set of similar vignettes (one improving, one worsening, two ambiguous) assessing reliability of the model;
2. clinicians were offered a POCT and a PCT test (they could select either, neither, or both) with the two providing conflicting results (POCT negative, PCT positive). This allowed us to assess whether there is a systematic

preference for either POCT or PCT, and how clinicians consolidate disparity between the two in their decision making;

3. clinicians had the option to stop, continue, escalate or de-escalate antibiotics.

We hypothesised that a negative POCT result would increase clinicians' inclination to stop antibiotics, as in the previous study (*hypothesis 1*) [16]. We also expected that this effect would be smaller than that observed in the previous study (*hypothesis 2*), because:

1. the presence of a borderline positive PCT result would reduce clinicians' inclination to stop, because clinicians may have greater trust in the (more established) PCT test than the (less established) PCR-POCT, or because its implication of the possible presence of infection may mitigate the negative POCT result;
2. allowing clinicians to (de)escalate antibiotics would reduce the incidence of stopping, as de-escalation may be perceived as a "safer" (i.e., less risky) alternative.

METHODS

Participants

Consultants and trainees from the ICU (with 3+ months continuous experience in ICU) working in London-based university hospitals were invited through advertisements in closed social media groups exclusive to ICU consultants and trainees. Clinicians who took part in the previous study [16] and had indicated a willingness to be involved in future studies, were contacted via direct email. The

email contained a direct link to the online survey hosted by Qualtrics (Washington, USA). The survey remained open from March – October 2022.

Sample size

A minimum of 77 responses were required to demonstrate the same effect size of an increase in antibiotic stop decisions from 54-70% [16] with 80% power at the alpha = 0.05 level. Further details are available in the Supplemental Materials (SM1).

Importantly, however, we expected the effect to be smaller in the present study than the previous one (see *hypothesis 2*). We therefore conducted a second sample size calculation, identical to the first except that it aimed to detect a smaller effect (w =0.20). After adjusting for clustered data, the number of responses required was 233 and the number of participants required was 58 (233/4). We therefore aimed to recruit 58 participants.

Materials

We constructed four clinical scenarios of resolving lung infection after a course of antibiotics. Each scenario comprised clinical and biological data, which were varied to create four distinct patient trajectories (Table 1).

Table 1. Clinical vignettes used in this study. Full details of the scenarios can be found in the Supplemental Materials (SM2).

Vignette name	**Description**
Improvement	A man in his 60s with lobar pneumonia. Clinico-biological improvement after a 5-day course of antibiotics.

Worsening	A man in his 70s with lobar pneumonia who deteriorates to requiring mechanical ventilation. After initial stabilisation after 5 days of antibiotics, there is a decline in clinical and biological status.
Discordant: clinically better, labs worse ("disc clin better")	A man in his 60s with a severe lobar pneumonia requiring mechanical ventilation who improves clinically at 7 days and is extubated after an antibiotic course, but whose blood biomarkers are worse.
Discordant: clinically worse, labs better ("disc clin worse")	A man in his 50s with multilobar pneumonia who completes a course of antibiotics, is extubated but then deteriorates clinically despite improving blood biomarkers of infection.

These vignettes were thought to accurately represent the varying degrees of diagnostic (un)certainty commonly encountered in the critical care setting. Two of the scenarios (*improvement* and *worsening*) acted as 'positive' and 'negative' controls, in that the trajectory clearly supported stopping antibiotics (*improvement*) or continuation/escalation of antibiotics (*worsening*). The remaining two scenarios presented uncertainty with regards to antibiotic decision making and were used to explore the importance of the clinical vs. biological trajectories.

In each scenario, two test results were made available: POCT and PCT. Both tests were described as valid and reliable. The POCT was an infection-identifying PCR test that provided rapid diagnostics for bacteria. The POCT result (unbeknown to clinicians at the point of request) was always negative (suggesting no active infection within the lung), while the PCT result was always marginally positive (suggesting possible infection). Notably, the POCT in these scenarios was an indicator of the actual presence of an infectious organism, whereas the PCT test was a surrogate marker for the presence of infection. The conflicting results of these two tests would

allow us to compare how a direct vs. indirect test for infection might influence clinicians' choices.

Procedure

Clinicians provided informed written consent before being able to proceed further in the online survey. Following this, they responded to the four vignettes, presented in a random order. Each vignette began with a brief patient description, including the patient's age, sex, details of admission and clinical status one week later (i.e., after completing a course of antibiotics). Based on this, clinicians were asked to decide the best course of action with regards to antibiotics (*stop antibiotics* / *continue with current course of antibiotics* / *escalate antibiotics* / *de-escalate antibiotics*). They were also asked to rate their confidence in this decision, on a 6-point Likert scale anchored at 1=*not at all confident* and 6=*extremely confident*.

Clinicians were then informed of the availability of a POCT and a PCT test. Clinicians could select the test/s that they wished to perform (*POCT only / PCT only / both POCT and PCT / neither POCT nor PCT*). They were also asked to indicate the reason(s) for this decision (SM3).

Clinicians that chose to perform one of the two tests (*POCT only / PCT only*) were at this point presented with the result of the chosen test (POCT was always negative, PCT always positive) and asked whether they would change their previous antibiotics decision (*yes / no*). Those who responded *yes* were offered the previous antibiotics options again (*stop / original course / escalate (more than original course) / de-escalate (less than original course)*). Regardless of whether their decision had changed, they were asked to indicate their confidence in their decision (1-6, as above) and to explain their decision (free text response). Following this, the second

(non-requested) test result was displayed (POCT was always negative / PCT was always positive) and clinicians were asked the same questions again; that is, they were asked whether they would change their decision (if *yes*, they were presented with the four antibiotics options described above), to indicate their confidence in their decision (1-6), and to explain their decision (free text).

Clinicians that chose to perform both or neither of the tests (*both POCT and PCT / neither POCT nor PCT)* were shown the results of both tests simultaneously (POCT negative, PCT positive) and asked exactly the same questions; i.e., whether they would change their decision (*yes / no*) and what they would change it to if so (*stop / original course / escalate (more than original course) / de-escalate (less than original course)*), confidence in decision (1-6) and rationale for decision (free text).

Importantly, therefore, clinicians always gave an "initial" antibiotic decision (no test results seen) and a "final" antibiotic decision (both test results seen). When clinicians chose to perform one of the two tests, they also gave "interim" decisions (that one test result seen).

After completing all four scenarios, clinicians were also asked to complete Grol et al.'s *Attitudes to Risk-Taking in Medical Decision Making* [17] questionnaire (adapted to the ICU setting) and to provide demographic information (gender and level of training). The study procedure is presented graphically in the Supplemental Materials (SM4). Details of the piloting processes are presented in the Supplemental Materials (SM5).

Statistical analysis

To measure the combined effect of the test results (negative POCT and positive PCT) on clinicians' inclination to stop antibiotics, we compared the proportion of clinicians that chose to stop antibiotics initially vs. finally, using chi-square analysis. We had intended to also measure and compare the effects of each test result by analysing the interim decisions of those who selected *POCT only* or *procalcitonin only*; however, very few participants selected these options (see Results), rendering such an analysis unreliable.

To explore this further, we created a continuous variable termed "willingness-to-stop", by signing confidence ratings in accord with antibiotic decisions. Specifically, confidence ratings (1=*not at all confident* to 6=*extremely confident*) were signed positive (+) if clinicians chose to stop antibiotics, or negative (-) if they chose to continue (be it in the form of *escalation*, *de-escalation*, or *continuation of original course*). Initial and final willingness-to-stop antibiotics (both ranging from -6=lowest to 6=highest) were compared using Wilcoxon signed ranked tests (non-parametric, as the willingness-to-stop variables were not normally distributed, see Results).

We also explored whether clinicians might have a preference for POCT vs. PCT tests, by examining the proportion of participants that requested (vs. rejected) each test. We explored whether this might differ by scenario, using chi-square analysis.

Finally, we explored the effect of scenario (1=*improvement*, 2=*disc clin better/worse*,3=*worsening*), initial antibiotic decision (0=*escalate*, 1=*original course*, 2=*de-escalate*, 3=*stop*), test(s) requested (1=*requested* and 0=*rejected* for POCT and procalcitonin, respectively), *Attitude To Risk-Taking* score (the per-participant sum of responses to Grol et al.'s questionnaire), and level of experience (0=trainee, 1=consultant) on final antibiotic decisions (0=*escalate*, 1=*original course*, 2=*de-*

escalate, 3=*stop*), using a mixed effects ordinal logistic regression model with a per-participant random intercept.

A *p*-value < 0.05 was considered as statistically significant. Statistical analysis and graphing were performed using SPSS 28 (IBM, New York, USA) and Stata/MP 17 software (StataCorp, Texas, USA).

Approval

This study was approved by the Imperial College Research Ethics Committee (ICREC reference 20IC6499) and the manuscript adheres to CHERRIES guidance for reporting e-survey results [18].

Patient and public involvement

None.

RESULTS

Demographic data

Sixty six clinicians completed the survey. The number of clinicians that accessed the survey is unknown, as incomplete responses were deleted automatically, following one week of inactivity. All participants completed all four scenarios, providing a total of 264 scenario responses. Of the 66 clinicians, 39 (59.1%) were male. There were 34 (51.5%) consultants, 17 (25.8%) Specialst registrar, SpR clinicians, 13 (19.7%) Senior house officer, SHO clinicians and 2 (3%) Foundation Year, FY clinicians. The demographics of the sample can be found in the Supplemental Materials (SM6).

Initial and final antibiotic decisions

Figure 1 shows participants' initial and final antibiotic decisions. Prior to receiving any test results (POCT or PCT), participants opted to stop antibiotics 41.7% of the time (110/264). This was reduced to 35.6% (94/264) following the negative POCT and positive PCT test. Therefore, clinicians were less likely (rather than more likely) to stop antibiotics after receiving the test results (chi-square=4.03, df =1, p=0.045). Few clinicians chose the path of de-escalation initially or subsequent to the POCT/PCR results, with most opting to stop or escalate.

The frequency of stop decisions within each scenario is displayed in Figure 2. Within the *improvement*, *disc clin better* and *disc clin worse* scenarios, fewer participants chose to stop antibiotics in their final decision (following both test results) compared to their initial decision. This effect was highly significant in the *improvement* scenario (72.7% vs. 86.4%; chi square=10.50, df=1, p=0.001), significant in *disc clin better* (28.8% vs. 42.4%; chi-square=5.01, df=1, p=0.025), and non-significant in *disc clin worse* (36.4% vs. 37.9%, chi-square=0.07, df=1, p=0.797). In the *worsening* scenario, no participants chose to stop antibiotics initially, and was statistically unchanged following the test results (4.5%, p=*ns*). For a full breakdown of initial and final antibiotic decisions per scenario (i.e., number of participants that elected to *stop*, *continue with the original course*, *escalate*, and *de-escalate*), see the Supplemental Material (SM7).

Initial and final willingness-to-stop were nonparametric following inspection of frequency distribution histograms. This was confirmed using Kolmogorov-Smirnov tests (p<0.001 for both initial and final willingness-to-stop). Figure 3 shows clinicians'

median willingness-to-stop before and after receiving the negative POCT and positive PCT results, per scenario. Similar to Figure 2, willingness-to-stop appeared to decrease from initial to final decision in the *improvement*, *disc clin better* and *disc clin worse* scenarios; the effect was significant in the *improvement* scenario (z=-3.84, p<0.001, effect size r=0.33), significant in the *disc clin better* scenario (z=-2.56, p=0.010, r=0.22), and non-significant in the *disc clin worse* scenario (z=-0.11, p=0.909, r=0.01). In the *worsening* scenario, willingness-to-stop was greater in the final decision than the initial decision, suggesting that the negative POCT might increase clinicians' willingness-to-stop in this specific scenario. This effect was significant (z=-3.04, p=0.002, r=0.26). In Figure 3, the wide boxes in the *disc clin better* and *disc clin worse* scenarios suggest a greater variation in willingness-to-stop, as opposed to the narrower boxes in the *improvement* and *worsening* scenarios. *The* evolution of the willingness to stop from the initial decision, through an interim option and then the final decision is shown in the Supplemental Material (SM8).

POCT and PCT requests

Across all scenarios, POCT was requested 53.4% of the time (141/264), however this varied significantly by scenario (chi-square=55.97, df=3, p<0.001; *improvement=15.2%, disc clin better=57.6%, disc clin worse=68.2%, worsening=72.7%*). Similarly, PCT was requested 61% of the time (161/264) and also varied significantly by scenario (chi-square=52.01, df=3, p<0.001; *improvement=24.2%, disc clin better=74.2%, disc clin worse=78.8%, worsening=66.7%*).

Figure 4 displays the tests requested by clinicians per scenario. In the *improvement* scenario, the majority of participants requested neither the POCT nor PCT test (72.7%), potentially due to the patient's unambiguously positive trajectory (ceiling effect). Within the *disc clin better, disc clin worse* and *worsening* scenarios, most participants requested both tests (53.0%, 60.6% and 60.6% respectively). Participants' reasons for requesting/rejecting the tests were mainly related to supplementing clinical judgement and deeming tests necessary or not (SM9).

Factors influencing the final decision

Our mixed effects model explored the effect of 1) initial antibiotic decisions, 2) test/s requested, 3) attitudes toward risk taking, 4) levels of experience, and 5) scenarios on final antibiotic decisions. The results showed that initial antibiotic decisions had a strong effect upon final antibiotic decisions (OR=8.70 [95% confidence interval 5.40 to 14.02], p<0.001). That is, clinicians who were more [less] inclined to stop antibiotics prior to receiving the test results were also more [less] inclined to stop after receiving them. This was consistent with the original study (16). Additionally, participants that requested PCT (either alone or in conjunction with POCT) were less inclined to stop in their final antibiotics decision (OR=0.26 [0.12 to 0.55], p<0.001), whereas participants that requested POCT (either alone or in conjunction with PCT) were more inclined to stop in their final antibiotics decision (OR=3.84 [1.61 to 9.15], p=0.002). This was consistent with the original study [16]. Level of experience (trainee or consultant) had no significant effect on final antibiotics decisions (OR=1.56 [0.85 to 2.86], p=0.156), nor did differences in attitudes towards risk-taking (OR=1.02 [0.92 to 1.12], p=0.761) or scenario (OR=0.63 [0.32 to 1.22], p=0.168).

These findings did not change when we replaced initial and final antibiotic decisions with initial and final willingness-to-stop (SM10).

DISCUSSION

Antibiotic stop decisions did not increase following a negative PCR POCT result and borderline positive PCT result, after a completion course of antibiotics in ICU related respiratory infection. Rather, in three out of four scenarios (*improvement*, *disc clin better* and *disc clin worse*), the stop rate decreased. This differed from the first WHYSTOP study [16], where antibiotic stop decisions increased consistently (i.e., in all scenarios) following receipt of a negative POCT. Thus, a negative POCT did not increase stop decisions when there was a borderline positive PCT result.

A few potential explanations for the lack of increase in antibiotic stop decisions are likely. First, the addition of a borderline positive PCT result likely negated the effect of the negative POCT and so reduced clinicians' inclination to stop. Second, participants may have been more inclined to trust the PCT result (vs newer non established PCR POCT) due to its known potential and growing evidence base within clinical medicine [4,6,7]. Alternatively, participants may have had equal trust in PCT and POCT, but conflicting test results triggered a conservative approach to 'err on the side of caution' (i.e., continue antibiotics) [10]. This was likely, as demonstrated in the original WHYSTOP study. Specifically, in that study, the improvement scenario was extended to include the following 'twist': the original negative PCR POCT result was declared erroneous (due to lab error) and re-testing gave a positive result. In light of this, clinicians were given the opportunity to revise their antibiotic decisions. The proportion of antibiotic stop decisions, even in this clinico-biological case of improvement

(suggesting the resolution of infection), fell from 90% to 61% [16]. The caution adopted relates to 'prospect theory'; the idea that a loss and regret is psychologically twice as impactful as a gain; the context being potential patient harm from an antibiotic stop decision [19,20]. Another explanation may be that the clin better, biologically worse scenario may have been slightly more ambiguous in its clinical trajectory than in the original paper – that may have accounted for less confidence in choosing to stop antibiotics post POCT.

Finally, we expected that the addition of a de-escalation option might reduce the incidence of stopping antibiotics (by providing a "less risky" alternative) and thus weaken the effect of the negative POCT. Contrary to expectation, the de-escalation option was rarely used (8.1% of initial decisions and 9.4% of final decisions) and is therefore unlikely to account for the present findings.

Participants who actively requested (vs. passively received) the POCT result were more inclined to stop antibiotics, whereas those who actively requested (vs. passively received) the PCT result were less inclined to stop. Clearly, interpretation and/or weighting of test results does not take place in isolation, but is dependent upon the perceived relevance, and/or value of the test in the decision making. This suggests that 'forcing' POCT on clinicians is unlikely to bring about effective change. Rather increasing awareness and trust in POCT may allow for greater confidence in its use and increase uptake. Indeed, Dhesi et al. identified lack of clinician trust as a potential barrier to POCT adoption in practice, which must be addressed if POCT is to be of value in ICU settings [21]. This was also a potential reason for the lack of improvement in patient outcomes in a quasi-randomised controlled trial, where use of PCR POCT was compared with routine laboratory PCR in guiding clinical management [22].

Clinicians' grade (consultant vs. trainee) did not influence the inclination to stop antibiotics, nor did clinicians' attitude towards risk taking. The WHYSTOP study was underpowered to address this [16]. However, in a subsequent study, a difference in clinical decision making was noted between novices (i.e. clinical medical students) and clinicians [23]. This may be of particular interest, given the potential for learnt behaviour on antibiotic decision making [24,25]. Indeed, medical students, tested in the same conditions of the original WHYSTOP study, were initially more conservative than ICU clinicians in STOP decisions. More students chose the POCT when offered than clinicians, and the negative result increased the proportion of stop decisions in all scenarios, to the same level as clinicians [23].

We made changes to the WHYSTOP study protocol to better reflect reality. These included new clinical vignettes with the same 4 trajectories, introduction of an additional optional PCT test and (de)escalation opportunities. Within these new, more realistic scenarios, a negative POCT, alongside a positive PCT, did not increase antibiotic stop decisions. This study has demonstrated that contrasting test results and other behavioural factors may impair the utility of POCT. This may explain the lack of effect of POCT in improving patient outcomes (including length of hospital stay and antibiotic free days) seen in some trials investigating the use of other point-of-care biomarkers [11,22]. Other studies have demonstrated the potential for PCR POCT in critical care. The INHALE WP1 study investigated two PCR-based tests for the diagnosis of pneumonia, with both tests proving to be more sensitive than routine microbiology [8]. As new PCR POCTs are emerging and developing, there is a clear need for further investigation of the situations where use of POCT for infection, should be implemented [26].

Finally, the diminished effect of the negative PCR POCT in these scenarios may simply have been due to the reintroduction of further ambiguity (i.e. the borderline positive PCT). Whether ambiguity is present in the clinico-biological setting (i.e. the 2 discordant scenarios), or introduced through competing signals of the test results, it is anticipated that the requested POCT measure will act as a 'final arbiter', when diagnostic uncertainty remains. In the higher risk setting of ICU infection decisions, most clinicians adopt caution [27]. To provide clarity then, it may be sensible to choose just one rather than both point of care tests as this 'final' arbiter.

Strengths and limitations

This study offered realistic conditions and options to clinicians when making their antibiotic stop decisions. Specifically, it explored clinicians' preference for POCT vs. PCT (if any), in four different clinico-biological trajectories, whilst enabling (de)escalation options. Furthermore, we made the PCT result borderline positive, so as to 1) better represent the complexities and uncertainties of clinical practice and 2) stress-test the effect of a negative PCR POCT (i.e., assess its influence in the face of conflicting biomarker information).

We were unable to reliably analyse interim decisions, due to the small number of participants that requested either POCT or PCT. A larger sample size may have increased the size of these subgroups, allowing us to isolate and compare the respective effects of POCT and PCT on antibiotic decision making. Another way to compare their effects would be to vary their results (positive vs. negative) systematically. Presently, the effects of POCT and PCT are intertwined, which limits the conclusions that we can draw. While these conclusions are interesting and

informative, further work is needed to gauge any hierarchical influence of POCT, and PCT. A direct comparison of POCT vs. PCT (with both being available but mutually exclusive) may help to answer other desired but unanswered questions; which (if any) is preferred between a direct or indirect biomarker of infection and which (if any) has a greater effect on final antibiotic decisions?

The study was sufficiently powered to detect a modest increase in antibiotic stop decisions before versus after POCT/PCT results, across scenarios. All other statistical tests and sub-analyses were exploratory and should be interpreted with caution. We did not adjust the significance threshold for multiple comparisons, but our findings were frequently robust ($p < 0.001$) and would likely remain significant even with such an adjustment, which supports their validity. Generalizability may have been limited by the recruitment of clinicians from mainly Academic training programmes. The influence of "system-level noise" cannot also be underestimated. "Noise" refers to variation within decision making and here with regards to clinical practice [28]. At a system level, variation between academic centres and hospitals (with different antibiotic prescribing policies and in-house training) can lead to noise. For example, clinicians in one hospital may be more [less] familiar with a given test than clinicians in another hospital, due to higher [lower] utilisation of that test in their hospital.

CONCLUSION

A negative POCT result does not appear to increase clinicians' inclination to stop antibiotics, when presented alongside a borderline positive PCT test. Conflicting test results could thus be one reason why POCT has failed to increase antibiotic free days

in the ICU [11,22]. Further research investigating the behavioural and trajectorial factors that might compete with or override POCT in decision making, alongside initiatives to increase clinicians' confidence in POCT, are imperative to improve its utility in the ICU. Ultimately, recognising the uncertainty in prescribing decision making, how it affects clinicians, and developing decision-making tools to support them in avoiding overreliance on antibiotics should be a future endeavour to improve antibiotic stewardship [29].

CONTRIBUTORS

SS conceived the idea. Guarantor is SS. SS, MN and TL developed the protocol. MN and TL created the data collection tool. SS, MN and TL analysed the data. SS, MN and TL wrote the first and subsequent drafts. SS, MN, TL, AS, LM and NM reviewed the manuscript.

FUNDING

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

ACKNOWLEDGEMENTS

We would like to thank and acknowledge Dr Sonia Mason, Dr Marco Morosin and Dr Pascal Frederiks for their assistance in piloting the survey. In addition, we would like to express appreciation for all the clinicians, part of the WHYSTOP consortium, who participated in this study.

COMPETING INTERESTS

None declared.

DATA AVILABILITY STATEMENT

Data is available in a public, open access repository [30].

REFERENCES

1. Laxminarayan R, Duse A, Wattal C, et al. Antibiotic resistance—the need for global solutions. The Lancet Infectious Diseases. 2013; 13 (12): 1057-1098. doi: 10.1016/S1473-3099(13)70318-9.
2. Vincent J, Sakr Y, Singer M, et al. Prevalence and Outcomes of Infection Among Patients in Intensive Care Units in 2017. JAMA. 2020; 323 (15): 1478-1487. doi: 10.1001/jama.2020.2717.
3. Wirz Y, Meier MA, Bouadma L, et al. Effect of procalcitonin-guided antibiotic treatment on clinical outcomes in intensive care unit patients with infection and sepsis patients: a patient-level meta-analysis of randomized trials. Critical Care (London, England). 2018; 22 (1): 191. doi: 10.1186/s13054-018-2125-7.
4. Schuetz P, Beishuizen A, Broyles M, et al. Procalcitonin (PCT)-guided antibiotic stewardship: an international experts consensus on optimized clinical use. Clinical Chemistry and Laboratory Medicine. 2019; 57 (9): 1308-1318. doi: 10.1515/cclm-2018-1181.
5. Schuetz P, Christ-Crain M, Thomann R, et al. Effect of procalcitonin-based guidelines vs standard guidelines on antibiotic use in lower respiratory tract infections: the ProHOSP randomized controlled trial. JAMA. 2009; 302 (10): 1059-1066. doi: 10.1001/jama.2009.1297.
6. de Jong E, van Oers JA, Beishuizen A, et al. Efficacy and safety of procalcitonin guidance in reducing the duration of antibiotic treatment in critically ill patients: a randomised, controlled, open-label trial. The Lancet Infectious Diseases. 2016; 16 (7): 819-827. doi: 10.1016/S1473-3099(16)00053-0.
7. Bouadma L, Luyt C, Tubach F, et al. Use of procalcitonin to reduce patients' exposure to antibiotics in intensive care units (PRORATA trial): a multicentre randomised controlled trial. The Lancet. 2010; 375 (9713): 463-474. doi: 10.1016/S0140-6736(09)61879-1.
8. Enne VI, Aydin A, Baldan R, et al. Multicentre evaluation of two multiplex PCR platforms for the rapid microbiological investigation of nosocomial pneumonia in UK ICUs: the INHALE WP1 study. Thorax. 2022; doi: 10.1136/thoraxjnl-2021-216990.

9. Poole S, Clark TW. Rapid syndromic molecular testing in pneumonia: The current landscape and future potential. The Journal of Infection. 2020; 80 (1): 1-7. doi: 10.1016/j.jinf.2019.11.021.
10. Pandolfo AM, Horne R, Jani Y, et al. Understanding decisions about antibiotic prescribing in ICU: an application of the Necessity Concerns Framework. BMJ quality & safety. 2022; 31 (3): 199-210. doi: 10.1136/bmjqs-2020-012479.
11. Hellyer TP, McAuley DF, Walsh TS, et al. Biomarker-guided antibiotic stewardship in suspected ventilator-associated pneumonia (VAPrapid2): a randomised controlled trial and process evaluation. The Lancet Respiratory Medicine. 2020; 8 (2): 182-191. doi: 10.1016/S2213-2600(19)30367-4.
12. Kostopoulou O, Mousoulis C, Delaney B. Information search and information distortion in the diagnosis of an ambiguous presentation. Judgment and Decision Making. 2009; 4 (5): 408-418. doi: 10.1017/S1930297500001236
13. Kostopoulou O, Russo JE, Keenan G, et al. Information distortion in physicians' diagnostic judgments. Medical Decision Making: An International Journal of the Society for Medical Decision Making. 2012; 32 (6): 831-839. doi: 10.1177/0272989X12447241.
14. Croskerry P. Achieving Quality in Clinical Decision Making: Cognitive Strategies and Detection of Bias. Academic Emergency Medicine. 2002; 9 (11): 1184-1204. doi: 10.1197/aemj.9.11.1184.
15. Saposnik G, Redelmeier D, Ruff CC, et al. Cognitive biases associated with medical decisions: a systematic review. BMC medical informatics and decision making. 2016; 16 (1): 138. doi: 10.1186/s12911-016-0377-1.
16. Singh S, Nurek M, Mason S, et al. WHY STOP? A prospective observational vignette-based study to determine the cognitive-behavioural effects of rapid diagnostic PCR-based point-of-care test results on antibiotic cessation in ICU infections. BMJ Open. 2023 Nov 21;13(11):e073577. doi: 10.1136/bmjopen-2023-073577.
17. Grol R, Whitfield M, De Maeseneer J, et al. Attitudes to risk taking in medical decision making among British, Dutch and Belgian general practitioners. The British Journal of General Practice: The Journal of the Royal College of General Practitioners. 1990; 40 (333): 134-136. PMID: 2115347

18. Eysenbach G. Improving the Quality of Web Surveys: The Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res 2004; 6 (3): e34. DOI: 10.2196/jmir.6.3.e34
19. Kahneman, Daniel & Tversky, Amos, 1979. " Prospect Theory: An Analysis of Decision under Risk ," Econometrica, Econometric Society, vol. 47 (2), pages 263-291, March. Handle: RePEc:ecm:emetrp:v:47:y:1979:i:2:p:263-91
20. Zeelenberg M, van den Bos K, van Dijk E, et al. The inaction effect in the psychology of regret. J Pers Soc Psychol 2002;82:314–27
21. Dhesi Z, Enne VI, O'Grady J, et al. Rapid and Point-of-Care Testing in Respiratory Tract Infections: An Antibiotic Guardian? ACS pharmacology & translational science. 2020; 3 (3): 401-417. doi: 10.1021/acsptsci.0c00027.
22. Andrews D, Chetty Y, Cooper BS, et al. Multiplex PCR point of care testing versus routine, laboratory-based testing in the treatment of adults with respiratory tract infections: a quasi-randomised study assessing impact on length of stay and antimicrobial use. BMC Infectious Diseases. 2017; 17. doi: 10.1186/s12879-017-2784-z.
23. Singhal A, Nurek M, Lau T et al. Infection Point of Care Tests (POCT) in simulated vignettes can improve the education of Antibiotic Stewardship Programmes. An observational comparative study of Clinicians vs Medical Students. Research Square [Preprint] September 22, 2023. https://doi.org/10.21203/rs.3.rs-3295414/v1
24. Trimble M, Hamilton P. The thinking doctor: clinical decision making in contemporary medicine. Clinical Medicine. 2016; 16 (4): 343-346. doi: 10.7861/clinmedicine.16-4-343.
25. Croskerry P. Clinical cognition and diagnostic error: applications of a dual process model of reasoning. Advances in Health Sciences Education: Theory and Practice. 2009; 14 Suppl 1 27-35. doi: 10.1007/s10459-009-9182-2.
26. Newcombe V, Coats T, Dark P, et al. The future of acute and emergency care. Future Healthcare Journal. 2021; 8 (2): e230-e236. doi: 10.7861/fhj.2021-0097.
27. Pandolfo AM , Horne R , Jani Y, et al. Understanding decisions about antibiotic prescribing in ICU: an application of the necessity concerns framework. BMJ Qual Saf 2022;31:199–210. doi:10.1136/bmjqs-2020-012479

28. Kahneman D, Sibony O, Sunstein C. Noise: A Flaw in Human Judgement. London: William Collins; 2021.
29. Tarrant C, Krockow EM. Antibiotic overuse: managing uncertainty and mitigating against overtreatment. BMJ Qual Saf 2022;31:163–167. doi:10.1136/bmjqs-2021-013615

[dataset] 30. Lau T. Data from: Point of Care Tests, Diagnostic Uncertainty and Antimicrobial Stewardship in the ICU. Procalcitonin or PCR to Aid Antibiotic Stop Decisions; an Observational Cohort Study. OSF, October 9, 2024. doi:10.17605/OSF.IO/7GDES

FIGURE LEGENDS

Figure 1. Number and proportion of decisions made initially (before receiving any test results) vs. finally (after receiving the negative POCT and positive PCT results), regardless of scenario (n=264). Participants chose between the options of escalate (more than the original course), continue with the original course, de-escalate (less than the original course) and stop with regards to antibiotics.

Figure 2. Number of clinicians that chose to STOP antibiotics initially (i.e., before receiving any test results) vs. finally (after receiving both a negative POCT result and a positive PCT result), per scenario. The total number of responses in each scenario (improvement, disc clin better, disc clin worse, and worsening) was 66. * $p<0.05$

Figure 3. Median willingness-to-stop before vs. after receiving the negative POCT and positive PCT results, per scenario (n=66 for each scenario). Willingness-to-stop represents a participant's confidence (1-6), signed positive if they chose to stop antibiotics and negative if they chose to continue (be it via continuation of the original course, escalation, and de-escalation). * $p<0.05$

Figure 4. Number of participants requesting POCT and/or PCT tests, per scenario (n=66 for each scenario). Participants were offered a POCT and PCT test within each scenario, after making an initial antibiotics decision, and could request POCT only, PCT only, both tests or neither test.

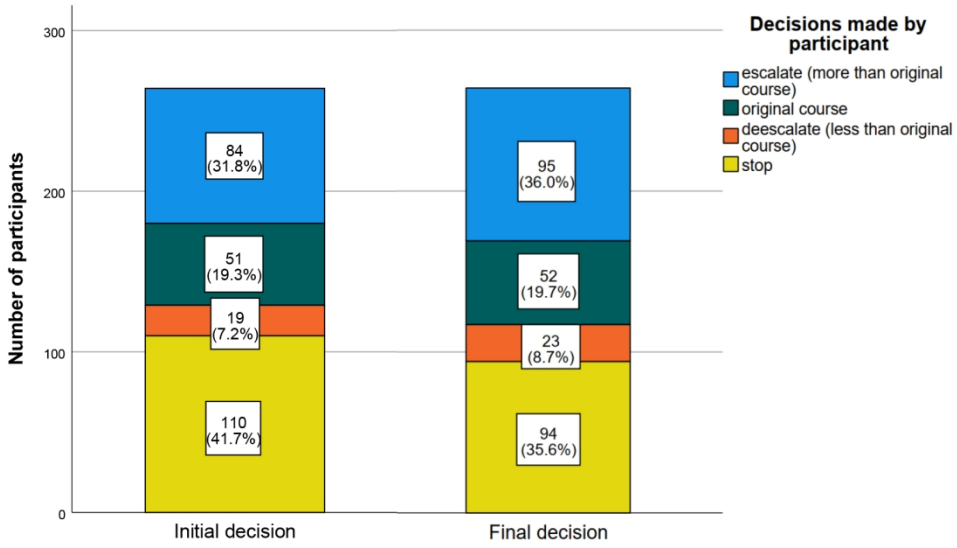


Figure 1. Number and proportion of decisions made initially (before receiving any test results) vs. finally (after receiving the negative POCT and positive PCT results), regardless of scenario (n=264). Participants chose between the options of escalate (more than the original course), continue with the original course, de-escalate (less than the original course) and stop with regards to antibiotics.

154x93mm (300 x 300 DPI)

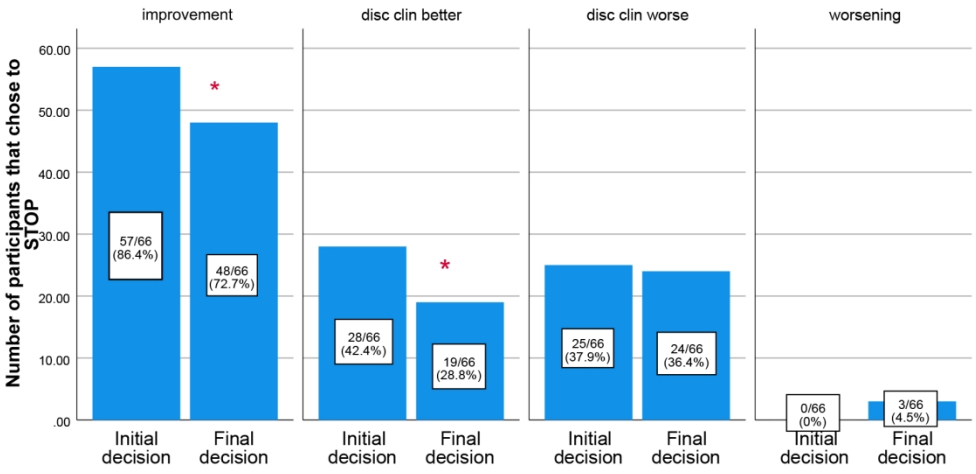


Figure 2. Number of clinicians that chose to STOP antibiotics initially (i.e., before receiving any test results) vs. finally (after receiving both a negative POCT result and a positive PCT result), per scenario. The total number of responses in each scenario (improvement, disc clin better, disc clin worse, and worsening) was 66. * p<0.05

159x93mm (300 x 300 DPI)

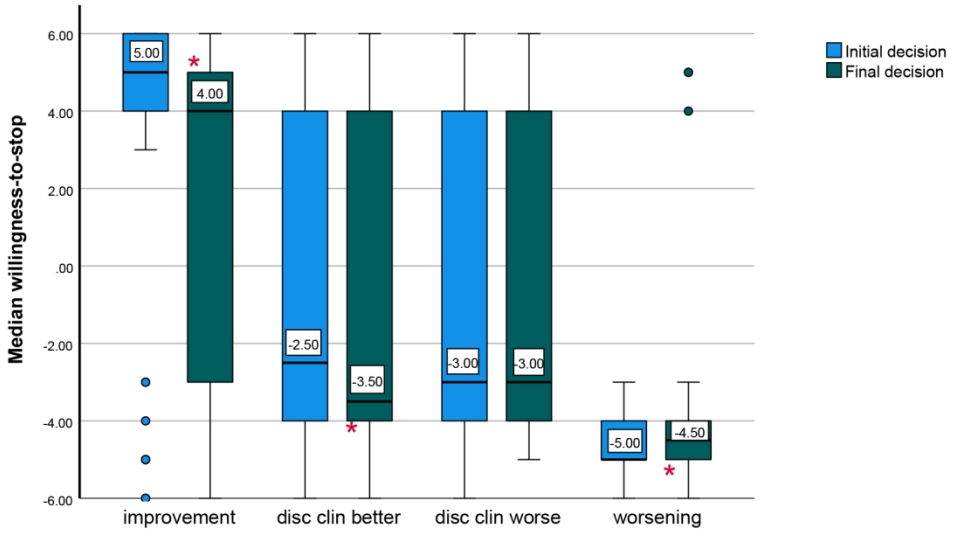


Figure 3. Median willingness-to-stop before vs. after receiving the negative POCT and positive PCT results, per scenario (n=66 for each scenario). Willingness-to-stop represents a participant's confidence (1-6), signed positive if they chose to stop antibiotics and negative if they chose to continue (be it via continuation of the original course, escalation, and de-escalation). * p<0.05

155x92mm (300 x 300 DPI)

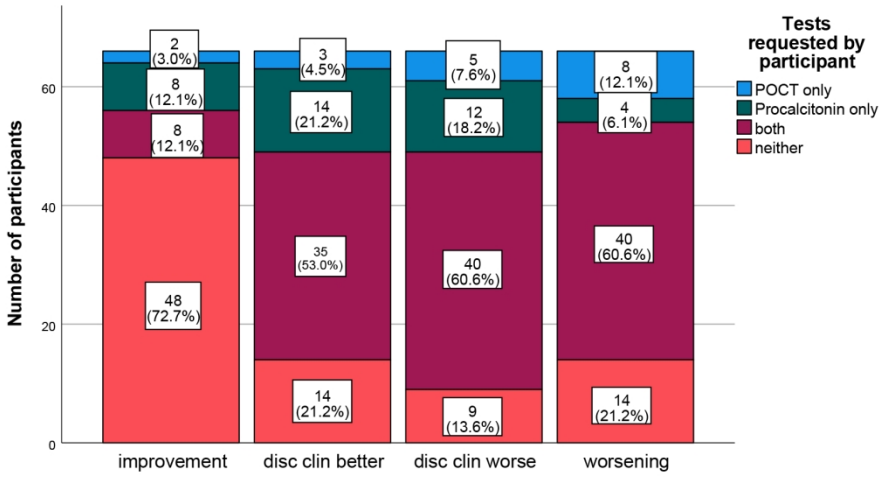


Figure 4. Number of participants requesting POCT and/or PCT tests, per scenario (n=66 for each scenario). Participants were offered a POCT and PCT test within each scenario, after making an initial antibiotics decision, and could request POCT only, PCT only, both tests or neither test.

159x93mm (300 x 300 DPI)

SUPPLEMENTAL MATERIAL

SM1

Sample size calculation.

The original WHYSTOP study found that a negative POCT result increased clinicians' willingness to stop antibiotics significantly ($p<0.01$). Specifically: prior to receiving the negative POCT result, clinicians were willing to stop antibiotics 54% of the time (138/258); after receiving the result, they were willing to stop antibiotics 70% of the time (180/258; chi-square=25.82, df=1, $p<0.01$, $w=0.32$). Using G*Power 3.1, we estimated that a minimum of 77 responses would be required to replicate this effect, with power at 80%, alpha at 0.05, and 1 degree of freedom. To account for clustered data (with each participant seeing 4 scenarios), we then calculated the "design effect" (DE)(1), using the formula $1+(n-1)\rho$, where n is the cluster size (4) and ρ the intraclass correlation/Cronbach's alpha (2) from Singh et al.'s study (0.061). Multiplying the number of required responses (77) by the DE (1.183) suggested that 91 responses were needed (77 x 1.183). At 4 responses per participant, 23 participants were required (91/4).

1. Donner A, Klar N. *Design and Analysis of Cluster Randomization Trials in Health Research.* London: Arnold; 2000. Available from: https://www.jameslindlibrary.org/donner-a-klar-n-2000/ [Accessed May 24, 2022].
2. Bi J, Kuesten C. Intraclass Correlation Coefficient (ICC): A Framework for Monitoring and Assessing Performance of Trained Sensory Panels and Panelists. *Journal of Sensory Studies.* 2012; 27 (5): 352-364. doi: 10.1111/j.1745-459X.2012.00399.x.

SM2a

The *improvement* vignette

Details of admission:

A 68-year-old male presented following a fall at home. He sustained rib fractures to his right anterior 5th and 6th ribs and was admitted for pain control. He has a background of poorly controlled insulin-dependent type 2 diabetes. Two days into admission, he developed hypoxia and pyrexia. His observations were the following:

Respiratory rate	26/min
SpO2	90% on room air
Heart rate	90/min sinus rhythm
Blood pressure	111/81 mmHg
Temperature	38.0 C

There was right sided consolidation on his chest radiograph and nil else. His blood tests demonstrated a WBC of 15 and CRP of 78. He was initiated on Co-amoxiclav.

Five days later:

7 days into his admission (5 days following antibiotics), he had improved shortness of breath and was afebrile. His observations were:

Respiratory rate	18/min
SpO2	99% on room air
Heart rate	83/min sinus rhythm
Blood pressure	112/80 mmHg
Temperature	37.0 C

He was pain free and mobilising on the ward. His repeat blood tests demonstrated a WBC of 8 and CRP of 15.

SM2b

The *overall worse* vignette

Details of admission:

A 78-year-old male was admitted with a 4-day history of worsening shortness of breath and a productive cough. He has a background of hypertension, Type II Diabetes Mellitus and a previous TIA (2019). He has no known drug allergies. His admission observations were:

Respiratory rate	22/min
SpO2	87% on room air
Heart rate	101/min sinus rhythm
Blood pressure	106/62 mmHg
Temperature	37.9 C

There was right basal consolidation on his chest radiograph. His blood tests demonstrated a WBC of 12 and a CRP of 70. He was empirically started on Levofloxacin and Clarithromycin. Within 24 hours he deteriorated and required mechanical ventilation.

Five days later:

5 days into his admission and after an initial improvement in ventilation, he became febrile. His observations were:

Respiratory rate	25/min
SpO2	90% on FiO2 21%
Heart rate	120/min sinus rhythm
Blood pressure	110/58 mmHg
Temperature	38.9 C

Further clinical assessment does not identify an alternative source of infection. However, his repeat blood tests demonstrated a WBC of 15 and a CRP of 150.

SM2c

The *disc clin better* vignette

Details of admission:

A 60-year-old male was admitted to ITU with a 6-day history of pyrexia, shortness of breath, and a productive cough with rusty sputum. He has a past medical history of well controlled, uncomplicated HIV (CD4 count >500 and viral load undetectable 3 months ago). His observations were the following:

Respiratory rate	32/min
SpO2	80% on room air
Heart rate	115/min sinus rhythm
Blood pressure	95/52 mmHg
Temperature	38.4 C

A chest radiograph demonstrated left midzone and basal consolidation. His blood tests demonstrated a WBC of 13 and a CRP of 50. Sputum culture grew MRSA. He was intubated and ventilated and empirically started on Linezolid. His initial blood gas findings were:

- FiO2 0.6
- PaO2 of 7.7 kPa
- PaCO2 5.2 kPa
- Base excess of -4

Seven days later:

7 days into his admission, he was improving on ventilation and extubated and weaned onto room air. He was feeling notably better. Clinical assessment does not identify any clinical source of infection. His observations were:

Respiratory rate	18/min
SpO2	96% on room air
Heart rate	82/min sinus rhythm
Blood pressure	124/70 mmHg
Temperature	36.9 C

However, his repeat blood tests demonstrated a worsening with a WBC of 15 and a CRP of 60.

SM2d

The *disc clin worse* vignette

Details of admission:

A 59-year-old male was admitted to ITU with a 3-day history of vomiting, pyrexia and a productive cough. He has a background of alcohol excess. His observations were the following:

Respiratory rate	30/min
SpO2	88% on room air
Heart rate	130/min sinus rhythm
Blood pressure	115/64 mmHg
Temperature	38.3 C

A chest radiograph demonstrated bilateral pulmonary infiltrates. His blood tests demonstrated a WBC of 20 and a CRP of 82. He was intubated and ventilated and empirically started on Piperacillin/Tazobactam. His initial blood gas findings on ventilation were:

- FiO2 0.5
- PaO2 of 8.2 kPa
- PaCO2 4.5 kPa
- Base excess of -5

Five days later:

5 days into his admission, he was making a good recovery, onto low pressure support ventilation and FiO2 down to 0.3. Chest radiograph findings were unchanged at this point. However, 1 day later, he developed new pyrexia to 37.8 C and increased oxygen requirement to FiO2 0.45. Investigations ruled out a pulmonary embolus.

Seven days later:

7 days into his admission, his observations were:

Respiratory rate	20/min
SpO2	92% on FiO2 0.45
Heart rate	100/min sinus rhythm
Blood pressure	130/70 mmHg
Temperature	37.8 C

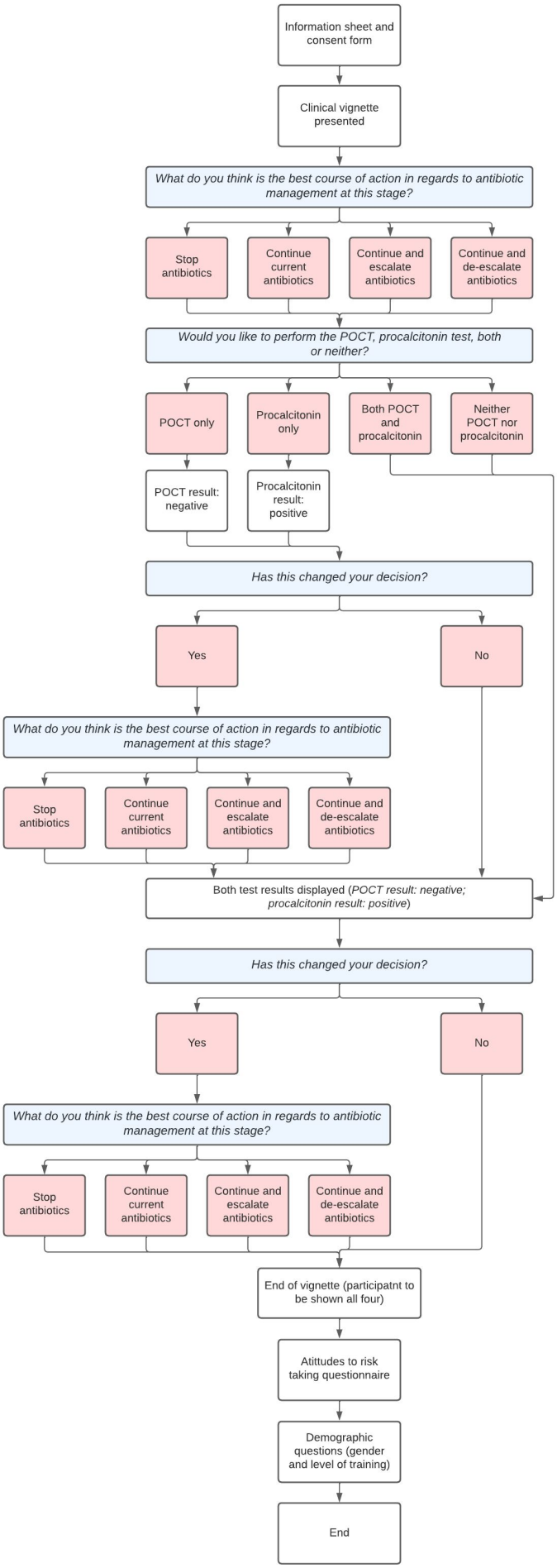
Clinical assessment does not identify an alternative source of infection.

His repeat blood tests demonstrated an ongoing reduction with a WBC of 10 and a CRP of 12.

SM3. Reasons for Clinicians' choice of diagnostic test when offered.

Reasons for performing POCT only	Reasons for performing PCT only	Reasons for performing both POCT and PCT	Reasons for performing neither POCT nor PCT

• I trust the POCT; • The POCT is necessary in this case; • I feel confident interpreting the POCT results; • I do not trust the procalcitonin test (I have concerns regarding the accuracy of the test); • The procalcitonin test is unnecessary in this case; • I do not feel confident interpreting the procalcitonin test results; • Other (if selected, the participant was asked to elaborate using free text).	• I trust the PCT test; • The PCT test is necessary in this case; • I feel confident interpreting the procalcitonin test results; • I do not trust the POCT (I have concerns regarding the accuracy of the test); • The POCT is unnecessary in this case; • I do not feel confident interpreting the POCT; • Other (if selected, the participant was asked to elaborate using free text).	• To supplement my clinical judgement; • I trust these tests; • The tests are necessary in this case; • I feel confident interpreting these tests; • Other (if selected, the participant was asked to elaborate using free text).	• I prefer to rely on my clinical judgement; • I do not trust these tests; • These tests are unnecessary in this case; • I don't feel confident interpreting these tests; • Other (if selected, the participant was asked to elaborate using free text).



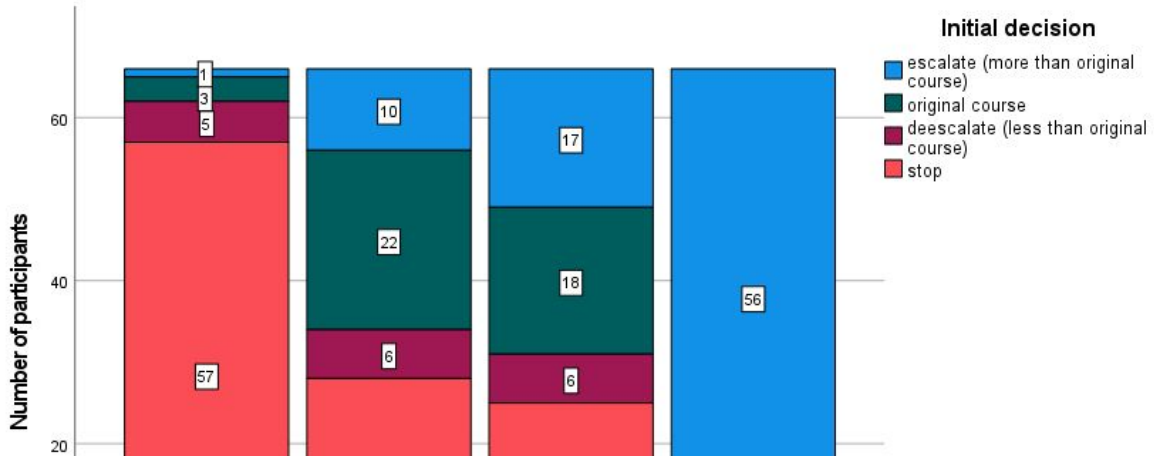
SM4. Graphical representation of the survey flow and procedure. Blue boxes indicate key questions, with the boxes below (red) displaying the possible responses. Created using Lucidchart (Lucid Software Inc., Utah, USA).

SM5. Piloting process

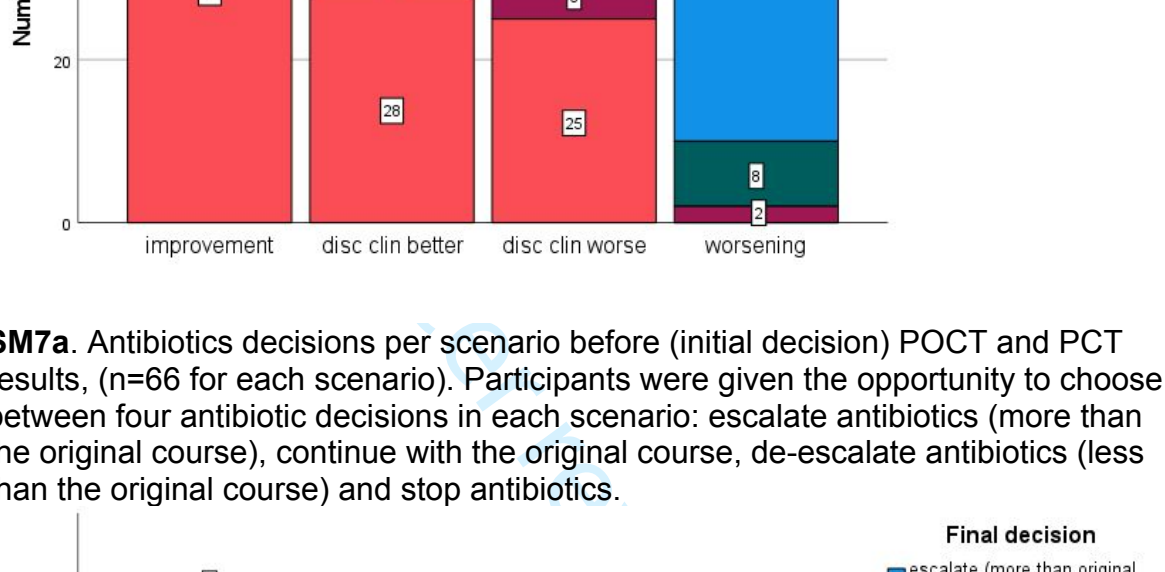
The survey was constructed and tested between the authors before piloting began. Two non-participating intensive care clinicians (SpR trainees) known to the authors were recruited to pilot-test the vignettes and survey. Feedback was given regarding the clarity and accessibility of the survey, as well as its format and structure. Feedback was very positive regarding the structure of the survey and contents of the vignettes, with only minor changes made to the survey. Particularly, we inserted a statement within the vignettes to suggest that there was no sign of alternative infection, as was suggested. Following this, the survey was trialed on other non participating ICU clinicians, then finalized and participant recruitment began.

SM6. Demographic and experience characteristics of the sample (n = 66).

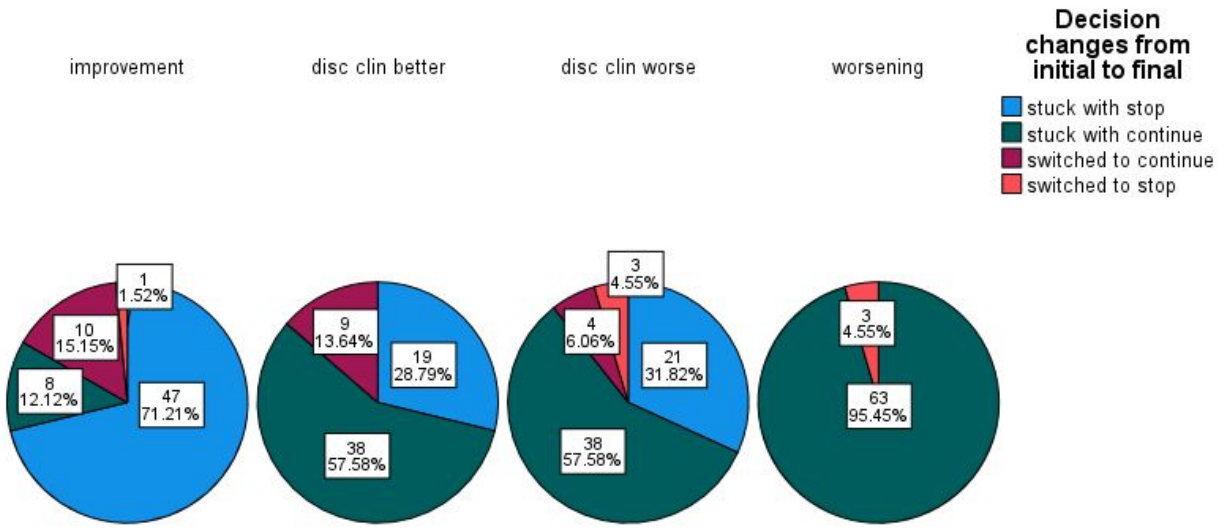
	n (%)	***Mean (SD), range***
Gender		
Male	39 (59.1%)	
Female	25 (37.9%)	
Prefer not to say	2 (3.0%)	
Grade		
Consultant	34 (51.5%)	
SpR trainee	17 (25.8%)	
SHO trainee	13 (19.7%)	
FY trainee	2 (3.0%)	
Experience: consultants		
Number of years since consultancy awarded		9.77 (7.67), 0 - 25
Experience: trainees		
>24 months on ICU ward	12 (37.5%)	
12-24 months on ICU ward	5 (15.6%)	
6-12 months on ICU ward	3 (9.4%)	
3-6 months on ICU ward	12 (37.5%)	



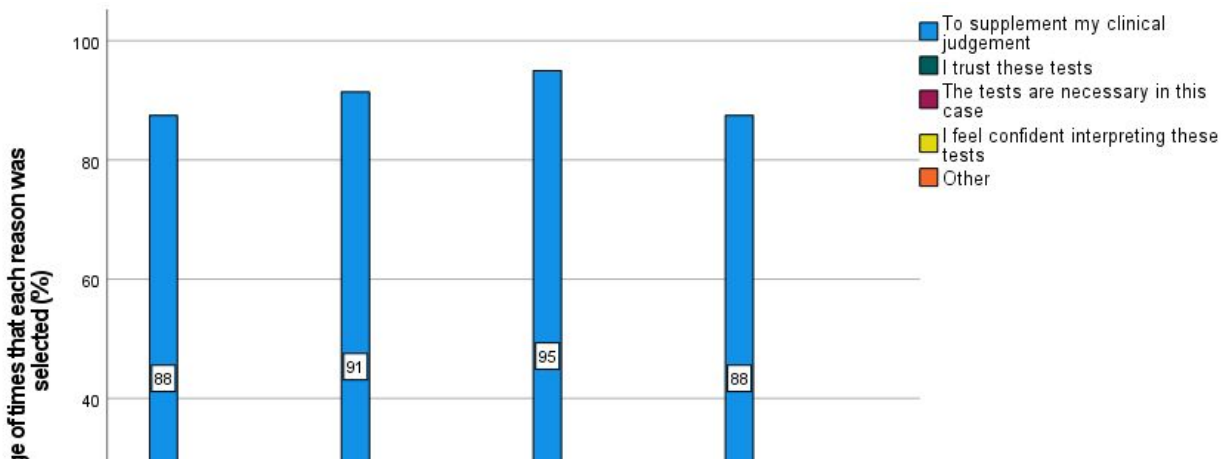
SM7a. Antibiotics decisions per scenario before (initial decision) POCT and PCT results, (n=66 for each scenario). Participants were given the opportunity to choose between four antibiotic decisions in each scenario: escalate antibiotics (more than the original course), continue with the original course, de-escalate antibiotics (less than the original course) and stop antibiotics.



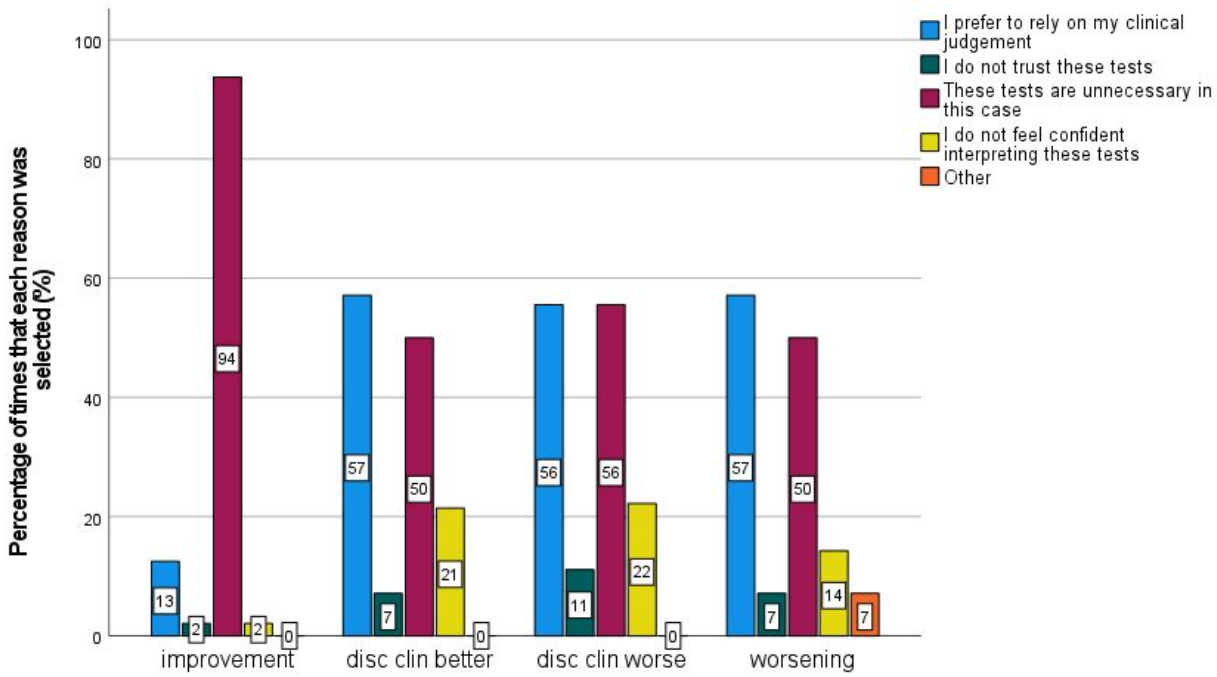
SM7b. Antibiotics decisions per scenario after POCT and PCT results (final decision), (n=66 for each scenario). Participants were given the opportunity to choose between four antibiotic decisions in each scenario: escalate antibiotics (more than the original course), continue with the original course, de-escalate antibiotics (less than the original course) and stop antibiotics.



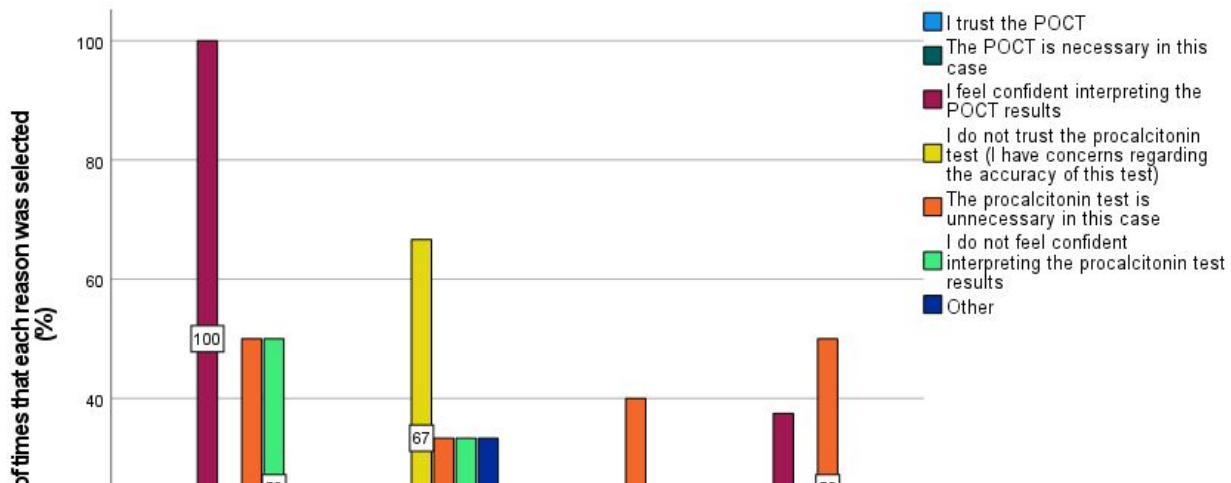
SM8. Decision changes, before vs. after receiving the negative POCT and positive procalcitonin test results, per scenario (n=66 for each scenario). Decisions were classified as “continue” if the participant elected to continue with the original course, escalate, or de-escalate. were all classified as “continue”. Absolute numbers and percentages are shown for each scenario.



SM9a. Reasons for decision to request **neither** POCT nor PCT given by clinicians in the improvement (n=48), disc clin better (n=14), disc clin worse (n=9) and worsening (n=14) scenarios. Participants were able to select more than one reason for their decision. Data labels indicate the percentage that each reason was selected within each scenario.



SM9b. Reasons for decision to request **both** the POCT and PCT given by clinicians in the improvement (n=8), disc clin better (n=35), disc clin worse (n=40) and worsening (n=40) scenarios. Participants were able to select more than one reason for their decision. Data labels indicate the percentage that each reason was selected within each scenario.



SM9c. Reasons for decision to request the **POCT only** given by clinicians in the improvement (n=2), disc clin better (n=3), disc clin worse (n=5) and worsening (n=8) scenarios. Participants were able to select more than one reason for their decision. Data labels indicate the percentage that each reason was selected within each scenario.

SM9d. Reasons for decision to request the **PCT only** given by clinicians in the improvement (n=8), disc clin better (n=14), disc clin worse (n=12) and worsening (n=4) scenarios. Participants were able to select more than one reason for their decision. Data labels indicate the percentage that each reason was selected within each scenario.

SM10. A mixed effects linear regression model of final willingness-to-stop on 1) initial willingness-to-stop, 2) test(s) requested, 3) attitudes towards risk taking, 4) level of experience (consultant vs. trainee), and 5) scenarios returned the following results. A p-value of less than 0.05 was considered as statistically significant.

Parameter	***b***	**95% Confidence Interval**	***p*-value**
Initial willingness-to-stop	0.70	[0.57 to 0.83]	<0.001
POCT request	1.41	[0.15 to 2.66]	0.028
PCT request	-1.54	[-2.71 to -0.38]	0.009
Experience	0.37	[-0.66 to 1.41]	0.483
Sum of attitude toward risk taking score	0.07	[-0.10 to 0.24]	0.418
Scenario	-0.32	[-1.10 to 0.46]	0.418
Constant	-0.90	[-3.41 to 1.61]	0.484

Item category	Checklist item	Page no.	Description
Design	Study design	2, 7	Prospective observational study. Simulated clinical scenarios.
	Ethics approval	13	Institutional Review Board approval.
	Informed consent	10	As part of the online Questionnaire. The survey began with an information sheet and consent form, which explained (inter alia) the length of time of the survey, which data were stored and where and for how long, who the investigators were and the purpose of the study. After the reading these, participants to provide informed consent by ticking a box labelled "I agree to participate in this study". If they did not tick this box, they were unable to proceed with the survey.
	Data protection		Personal information was not collected or stored as part of the survey. Participants were invited (but not required) to join our consortium ("WHY STOP"); those that elected to join were directed to a separate survey, where they were asked to provide their name and email address. These data were stored securely on a password-protected university computer. Survey responses remained wholly anonymous.
Development and pre-testing		11	This is explained in the manuscript methods and supplementary materials SM5. In an iterative process, surveys were repeatedly piloted and revised to ensure usability, technical functionality and plausibility of the clinical scenarios used.
Recruitment process	Open v Closed survey	7-8	The survey was accessible to consultants and trainees in Intensive Care Medicine (ICM; 3+ months continuous experience in ICU) currently working in London-based university teaching hospitals

	Contact mode	7-8	Initial contact was made via internet portals/forums and WhatsApp groups. Clinicians who took part in a previous survey and had indicated a willingness to be involved in future studies were contacted via direct email.
	Advertising the survey	7-8	The study was advertised via known trainee and consultant training portals (local), social media, and word of mouth.
Survey administration	Web/email	8	The survey was hosted on the secure Qualtrics platform and accessible via a website. Responses were downloaded from Qualtrics (i.e., captured automatically, not entered manually).
	Context	8	The website is the homepage of the WHY STOP Consortium, a diverse group of clinicians and behavioural scientists that aim to improve antibiotic stewardship in ICU (why-stop.wixsite.com/itu-decision-making). The website contains information about the Consortium, the core research team, and any past/present studies. For those who wish to participate in an open study, the website provides information on how to do so (e.g., a link to an open questionnaire). There was nothing on the website that might influence the selection of participants.
	Mandatory/voluntary	7-8	This was a voluntary study.
	Incentives	8	Study participants were invited to contribute to the manuscript, and to have this contribution acknowledged within the manuscript. This was entirely optional.
	Time/date	8	March-Oct 2022
	Item randomisation	10	Clinical vignettes were presented in a random order.
	Adaptive questioning	10-11	Clinicians were asked what tests they would like to request and could choose between POCT only, PCT only, both POCT and PCT, neither POCT nor PCT. They were then asked to

			indicate their reason/s by selecting from a list. These lists were different depending on the answer selected.
	Number of items	10-11	13 per vignette. At the end of the survey, participant's completed a measure of individual differences (5 items) and gave demographic information (5 items)
	Number of screens		25 in total (5 per patient case plus 5 for information sheet, consent form, introduction, risk inclination questionnaire and demographic questions).
	Completeness check		Mandatory items were enforced via Qualtrics' inbuilt functionality (participants could not proceed if they did not respond to mandatory items). Participants could select "prefer not to say" for demographic questions such as gender. Completeness was checked during piloting with junior doctors who were not involved in the actual survey.
	Review step		Participants were given the opportunity to update their answers during the course of the patient scenarios. This was an integral part of the study design.
Response rates	Unique site visitor		Each participant had a unique identifier as determined by the Qualtrics system ("ResponseID"). This was based on cookies (not IP Addresses).
	View rate		75 doctors started the survey. Of these, 66 completed all 4 scenarios
	Participation rate		75 doctors started the survey. Of these, 66 completed all 4 scenarios.
	Completion rate		75 doctors started the survey. Of these, 66 completed all 4 scenarios
Preventing multiple entries for same individual	Cookies used		When respondents started a session, Qualtrics placed a cookie on their browsers that kept track of their survey progress. Respondents had one week to return to the survey and finish their response. After a week, the response was recorded "as is". We made no attempt to prevent

			duplicate entries, given that clinicians had zero incentive (financial or otherwise) to do this and survey completion time was lengthy (20-25min).
	IP check		We made use of Qualtrics' "Anonymize responses" option, whereby client IP Addresses are not collected or stored. We did not check for duplicate responses, for the reasons stated above.
	Log file analysis		We did not check for duplicate responses, for the reasons stated above.
	Registration		The survey was accessed through an open link. We cannot be 100% sure that participants didn't fill in the form in duplicate. However, we think this highly unlikely given the length of the survey, the incentive structure (which did not change with submission of multiple surveys), and the degree of trust we maintain in colleagues.
Analysis	Handling of incomplete questionnaires	13	Incomplete responses were not utilised or analysed as they would be automatically removed after 1 week. Therefore, if a participant completed 3 scenarios and stopped mid-way through the 4th, these scenarios would not be used.
	Questionnaires with atypical timestamp		The minimum time needed to read and respond to a single scenario was 90 seconds (measured during piloting). We aimed to exclude respondents that took less time than this on a given scenario, but none did.
	Statistical correction	11-12	Not applicable. All statistical operations are described in the Analyses section of Methods.