Supplementary Appendix

Consensus position statement on advancing the classification of patients and tests of cure in studies of antibiotic treatment of complicated urinary tract infections

Authors: Truls E. Bjerklund Johansen, MD (Prof);[1] Christina Bahrs, MD;[2] Franck Bruyere, MD (Prof);[3] Tommaso Cai, MD;[4] Amelia Joseph, BMBS;[5] Béla Köves, MD;[6] David M. Livermore, PhD (Prof);[7] Alessandra Oliva, MD;[8] Alex Soriano, PhD;[9] Florian Wagenlehner, MD (Prof);[10] Per Henrik Zahl, PhD;[11] Manuel Etienne, MD (Prof);[12] cUTI consensus group

Affiliations:

1. Urology department, Institute of Clinical Medicine, University of Oslo, Oslo, Norway; and Institute of Clinical Medicine, University of Aarhus, Aarhus, Denmark
2. Division of Infectious Diseases, Department of Medicine I, Medical University of Vienna, Vienna, Austria; and Institute of Infectious Diseases and Infection Control, Jena University Hospital/Friedrich-Schiller-University, Jena, Germany
3. Urologie, CHRU Tours, Tours, France
4. Department of Urology, Santa Chiara Regional Hospital, Provincial Health Care Agency (APSS), Trento, Italy
5. Department of Clinical Microbiology, Nottingham University Hospitals NHS Trust, Queens Medical Centre, Nottingham, UK
6. Urology Department, South-Pest Teaching Hospital, Budapest, Hungary
7. Norwich Medical School, University of East Anglia, Norwich, UK
8. Department of Public Health and Infectious Diseases, Sapienza University of Rome, Rome, Italy
9. Department of Infectious Diseases, Hospital Clínic of Barcelona, University of Barcelona, IDIBAPS, CIBERINF, Ciber in infectious diseases ISCIII, Madrid, Spain
10. Department of Urology, Pediatric Urology and Andrology, Justus Liebig University Giessen, Giessen, Germany
11. Division for Mental and Physical Health, Norwegian Institute of Public Health, Oslo, Norway
12. Department of Infectious Diseases, CHU Rouen, Rouen, France; and INSERM UMR 1311 DYNAMICURE, Univ Rouen Normandie, Rouen, France.

Table of contents

Complicated urinary tract infection (cUTI) consensus group

In addition to the authors listed on the manuscript, the consensus group consisted of the following:

Name	Affiliation
Deepak Batura, FRCS	Urology Department, London North West University Healthcare NHS Trust, London, UK
Charlotte Graugaard-Jensen, MD	Department of Urology, Aarhus University Hospital, Aarhus, Denmark
Keith S. Kaye, MD (Prof)	Department of Medicine, Rutgers Robert Wood Johnson Medical School, New Brunswick, NJ, USA
Jennifer Kranz, FEBU	Department of Urology and Pediatric Urology, Uniklinik RWTH Aachen, Aachen, Germany; Department of Urology and Kidney Transplantation, Martin-Luther University, Halle, Germany
Massimiliano Lanzafame, MD	Unit of Infectious Diseases, Santa Chiara Hospital, Azienda Provinciale per i Servizi sanitari (APSS), Trento, Italy
José Medina-Polo, PhD	Department of Urology, Hospital Universitario 12 de Octubre-imas12, Madrid, Spain
Francisco Javier Membrillo de Novales, MD	CBRN and Infectious Diseases Department, Hospital Central de la Defensa "Gómez Ulla", Madrid, Spain
Kurt G. Naber, MD	Department of Urology, Technical University of Munich, Munich, Germany
J. Curtis Nickel, MD (Prof)	Department of Urology, Queens University and Kingston Health Sciences Center, Kingston, Ontario, Canada
Laila Schneidewind, FEBU	Inselspital, Dept. Of Urology, University Hospital Bern, University of Bern, Bern, Switzerland
Zafer Tandoğdu, FRCS (Urol)	Urology Department, University College London Hospitals, London, UK; Division of Surgery and Interventional Science, University College London, London, UK

Supplemental Methods

Healthcare professional tiering tool

A tiering tool is a digital algorithm used by public bodies and the pharmaceutical industry to identify health personnel with a certain clinical profile and expertise. Candidates are commonly classified into three levels based on formal qualifications, positions, and experience, and recognition among peers and the medical/scientific community. All members of the Delphi panel were considered to have the highest tier, being clinical scientists with an MD, PhD, or equivalent degree, and having proven expertise in urinary tract infections (UTIs).

Online platform

Within3 is a software communications company providing online discussion platforms for small or large groups of physicians and patients. The platform is developed using Ruby on Rails. User information is secured through a multilayered approach and certifications. Within3's primary and backup data centres are located at AWS, and the application maintains SOC2 Type 2 and ISO 27001:2013 certifications. Within3 has aligned controls with NIST 800-53 and GDPR guidelines and applicable cybersecurity legislation.

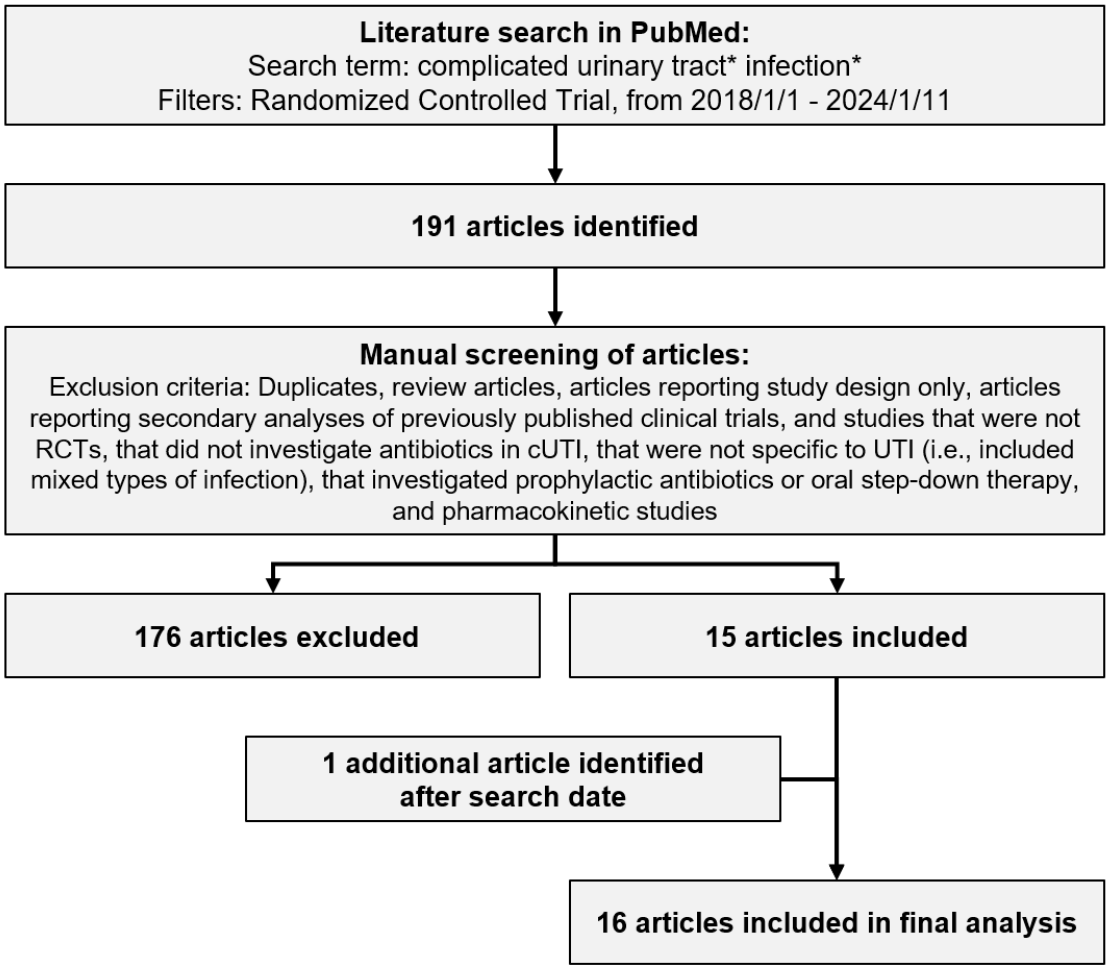
Literature search

PubMed was searched for articles with publication date from 1 January 2018 to 11 January 2024 for articles including the term: complicated urinary tract* infection*. Search results were filtered by article type to include only randomized controlled trials. Articles were manually screened to exclude duplicates, review articles, articles reporting study design only, articles reporting secondary analyses of previously published clinical trials, and studies that were not randomized controlled trials, that did not investigate antibiotics in cUTI, that were not specific to UTI (i.e., included mixed types of infection), that investigated prophylactic antibiotics or oral step-down therapy, and pharmacokinetic studies. The exclusion criteria were used to allow the most relevant papers to be identified. No studies were excluded based on publication language, but the final selected papers were all published in English. An additional important article published following completion of the literature search was added.[1] Finally, 16 papers were selected.

Literature analysis

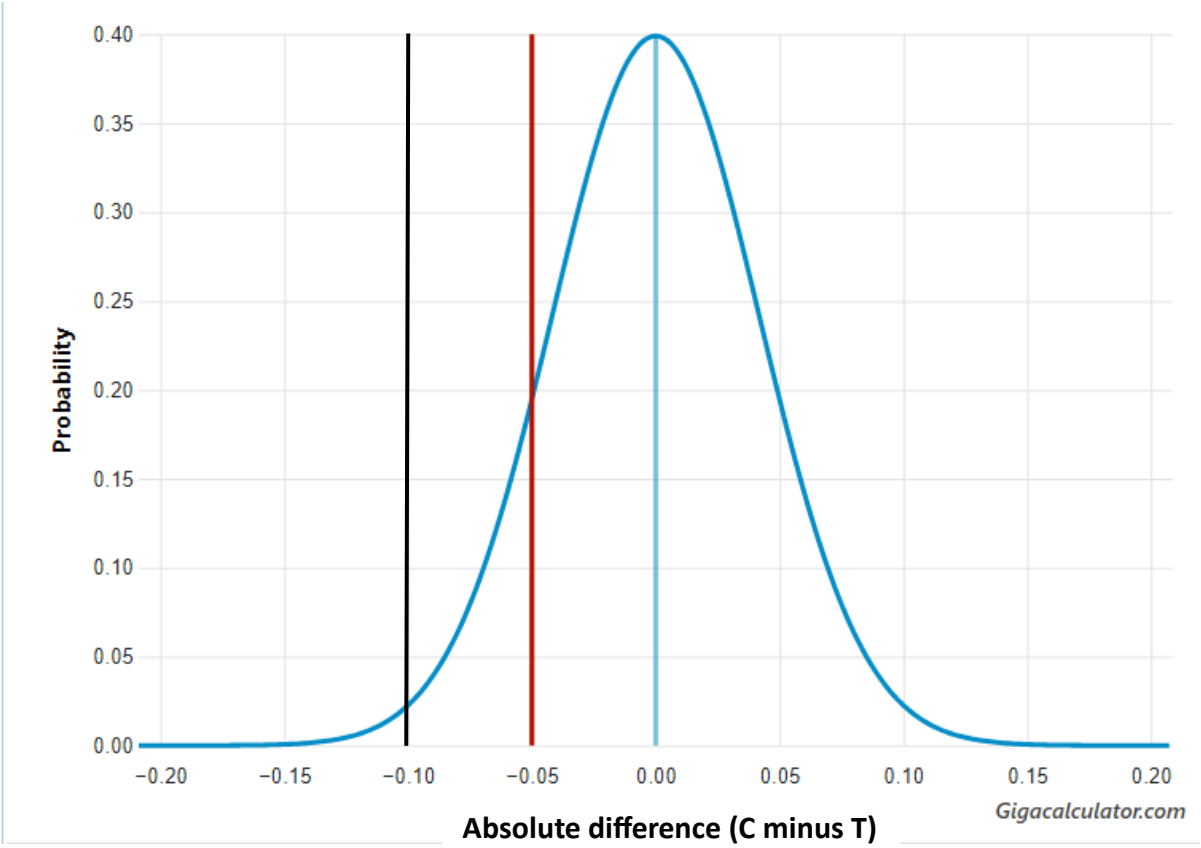
A spreadsheet with the US Food and Drug Administration (FDA) criteria for cUTI and risk factors defined by the group was developed, and the 16 papers identified by the literature search were evaluated for adherence to these criteria. The evaluation was descriptive in nature and the spreadsheet captured information about study protocols (e.g., study design, treatments, endpoints, eligibility criteria [including definition of cUTI and risk factors according to ORENUC criteria], timing of efficacy assessments), populations (e.g. characteristics reported for enrolled patients), pathogens (type and resistance), and treatment outcomes (e.g., superiority or non-inferiority assessments, discordant findings between clinical and microbiological outcomes).

Figure S1: Flowchart showing literature search and selection of studies for analysis



cUTI=complicated urinary tract infection. RCT=randomized controlled trial. UTI=urinary tract infection.

Figure S2: Statistical analysis – Interrelationship of parameters influencing the p-value of falsely rejecting a null hypothesis stating a difference between treatment arms in non-inferiority studies



Observed difference (red line), blue vertical line (no difference), dark blue line (error distribution), and non-inferiority criterium (black line). The curve is made by gigacalculator.com. The non-inferiority margin (black line with a value of –0·10) is derived through a combination of statistical analysis and clinical judgement.[2-4] It is a constant. In contrast, the observed treatment difference (red line) is a stochastic value. The p-value is the area to the right of the red line and under the error distribution. When the red line occurs on the left side of the non-inferiority margin, we use a corrected p-value, which is the area to the right of the black line and under the error distribution. The width of the error distribution (the distance between the two points where the second derivative equals zero, or the points where the slope changes from increasing to decreasing, or vice versa) depends on the number of individuals in the test and control groups. C=number of patients with treatment success in the control group. T=number of patients with treatment success in the test group (new antibiotic).

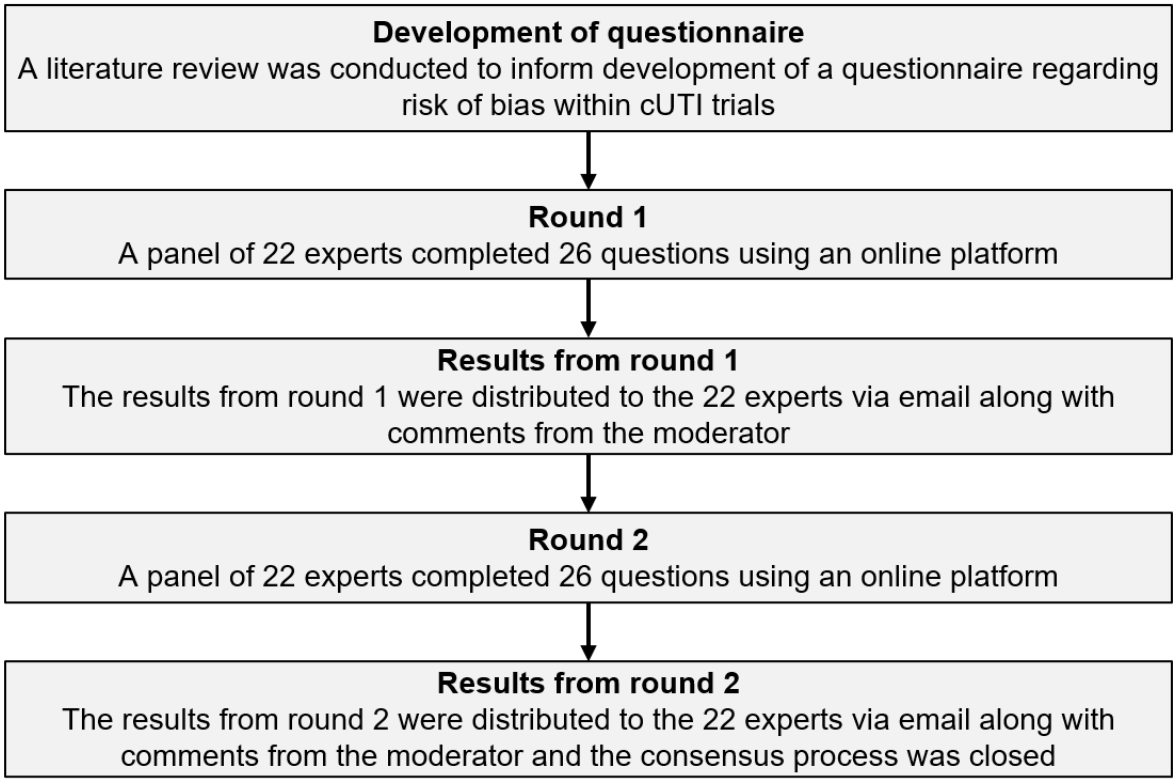
Table S1: Delphi survey

Signalling questions	Category	Issue number	Issues that are likely to affect the ability to draw reliable conclusions from studies	Risk of bias		
				Low	Some concerns	High
A. How would you assess the risk of bias in study outcomes in terms of clinical and microbiological success (including sustainability), if these issues (risk factors) were: • not considered at all, or • not equally distributed, or • not clearly defined (i.e., fever) in study groups?	Clinical situation	1	The type and frequency of clinical presentation and severity, i.e., percentages of cystitis, non-febrile and febrile pyelonephritis, and urosepsis			
		2	Whether the infection was nosocomial or community acquired			
		3	Fever			
		4	The general condition of the patient (ASA status)			
		5	If two or more of the issues in this category were not satisfactorily described			
	Patient	6	Catheter or stent (including nephrostomy tube) in place at diagnosis, during treatment and during follow-up			
		7	Presence of urinary stones anywhere in the urinary tract			
		8	History of symptomatic UTI in the previous 6 months			
		9	A history of bacterial prostatitis			
		10	Evidence of immune suppression or diabetes			
		11	Female sex			
		12	Antibiotic treatment within the previous 30 days			
		13	Premenopausal women who have a history of recurrent UTI associated with sexual intercourse			
		14	A history of obstipation			
		15	If two or more of the issues in this category were not satisfactorily described			
B. How would you assess the risk of bias of study outcomes in terms of microbiological success, if these issues related to the pathogen were: • not reported at all, or • not equally distributed in study groups?	Pathogen	16	The spectrum of pathogens			
		17	Resistance to study drugs			
		18	If both issues in this category were not satisfactorily described			
C. How would you assess the risk of bias when comparing studies if these differences in study characteristics were present?	Study characteristics	19	Difference between studies in the duration of treatment with study drug of more than 100% of average treatment duration (typically >3 days)			
		20	The period from EOT to TOC assessment differs by >1 week between studies			
		21	The observation period from EOT to assessment of sustainability of effect differs by >3 weeks between studies			
		22	A discrepancy between studies in microbiological success criteria of ≥10 CFU/mL (using $\leq 10^3$ instead of $\leq 10^4$ CFU/mL)			
		23	A discrepancy between studies in the definition of microbiological "eradication" (using $\leq 10^3$ CFU/mL or no definition at all)			
		24	If two or more of the issues in this category were present			

D. How would you assess the overall risk of bias in the outcomes of a study (in terms of clinical and microbiological success):	**All categories**	25	If at least one issue in each of the following categories was present (clinical situation, patient and pathogen)?			
E. How would you assess the overall risk of bias when comparing outcomes between studies (in terms of clinical and microbiological success):		26	If at least one issue in all four categories (clinical situation, patient, pathogen, and study characteristics) was present?			

For each question, participants used tick boxes to select the risk of bias and could also write free text comments. ASA=American Society of Anaesthesiologists. CFU=colony forming unit. EOT=end of treatment. TOC=test of cure. UTI=urinary tract infection.

Figure S3: Flowchart of modified, accelerated Delphi process



cUTI=complicated urinary tract infection.

Supplemental Results

Table S2: Overview of randomized controlled trials included in the literature analysis

Study	Phase	Blinding	Superiority and/or non-inferiority assessed?	Single/ multicentre (no. sites)	National/ multinational (no. countries)	Number of patients randomized	Primary endpoint	Test intervention	Comparator intervention
Bradley et al. 2019[5]	2	Single-blind	NA	Multi (25)	Multinational (9)	97	Safety and tolerability	Ceftazidime-avibactam IV	Cefepime IV
Connolly et al. 2018[6]	2	Double-blind	NA	Multi (27)	Multinational (4)	145	Microbiological eradication	Plazomicin 15 mg/kg or 10 mg/kg IV	Levofloxacin IV
Dunne et al. 2023[7]	3	Double-blind	Non-inferiority	Multi (131)	Multinational (13)	1395	Overall response (composite of clinical and microbiological success)	Sulopenem IV*	Ertapenem IV*
Eckburg et al. 2022[8]	3	Double-blind	Non-inferiority	Multi (95)	Multinational (15)	1372	Overall response (composite of clinical cure and microbiological response)	Tebipenem pivoxil hydrobromide (oral)	Ertapenem IV
Edlund et al. 2022[9]	4	Open-label	Non-inferiority and superiority	Multi (12)	National (1)	152	Disturbance of intestinal microbiota	Temocillin IV	Cefotaxime IV
Kaye et al. 2022[10]	3	Double-blind	Non-inferiority and superiority	Multi (90)	Multinational (19)	1041	Overall treatment success (composite of clinical cure and microbiological eradication)	Cefepime/enmetazobactam IV	Piperacillin/ tazobactam IV
Kaye et al. 2019[11]	2/3	Double-blind	Non-inferiority	Multi (92)	Multinational (16)	465	Overall success (composite of clinical cure and microbiological eradication)	ZTI-01 IV	Piperacillin/ tazobactam IV
Kaye et al. 2018[12]	3	Double-blind	Non-inferiority and superiority	Multi (60)	Multinational (17)	550	Co-primary endpoints: For FDA: Overall success (composite of clinical cure and microbiological eradication) For EMA: Microbiological eradication	Meropenem-vaborbactam IV	Piperacillin-tazobactam IV
Lafaurie et al. 2023[13,14]	3	Double-blind	Non-inferiority	Multi (27)	National (1)	240	Treatment success (composite of clinical success, microbiological success, and absence of a new antimicrobial treatment for UTI)	7 days of antibiotics†	14 days of antibiotics†
Leitner et al. 2021[15,16]	2/3	Double-blind‡	Non-inferiority and superiority	Single (1)	National (1)	113	Microbiological response	Intravesical pyophage	SoC antibiotics or intravesical placebo
Li et al. 2021[17]	3	Open-label	NA	Multi (34)	National (1)	208	Clinical cure	Sitafloxacin oral	Levofloxacin oral
Portsmouth et al. 2018[18]	2	Double-blind	Non-inferiority and post hoc superiority	Multi (67)	Multinational (15)	452	Composite of clinical response and microbiological response	Cefiderocol IV	Imipenem-cilastatin IV
Roilides et al. 2023[19]	2	Double-blind	NA	Multi (28)	Multinational (8)	134	Rates of adverse events and changes in laboratory values and vital signs	Ceftolozane tazobactam IV	Meropenem IV
Wagenlehner et al. 2024[1]	3	Double-blind	Non-inferiority and superiority	Multi (68)	Multinational (15)	661	Composite of both microbiologic and clinical success	Cefepime-taniborbactam IV	Meropenem IV
Wagenlehner et al. 2019[20]	3	Double-blind	Non-inferiority	Multi (68)	Multinational (14)	609	Composite cure (clinical cure and microbiologic eradication)	Plazomicin IV	Meropenem IV

Study	Phase	Blinding	Superiority and/or non-inferiority assessed?	Single/ multicentre (no. sites)	National/ multinational (no. countries)	Number of patients randomized	Primary endpoint	Test intervention	Comparator intervention
Wagenlehner et al. 2018[21]	2	Double-blind	NA	Multi (NS)	Multinational (2)	225	Combined clinical and microbiological response	Finafloxacin (IV then oral) for 10 days or 5 days total	Ciprofloxacin (IV then oral) for 10 days total

EMA=European Medicines Agency. FDA=US Food and Drug Administration. IM=intramuscular. IV=intravenous. NA=not assessed. no.=number. NS=not specified. UTI=urinary tract infection. *Sulopenem IV followed by oral sulopenem etzadroxil/probenecid or ertapenem IV followed by oral ciprofloxacin or amoxicillin-clavulanate. †Empirical ofloxacin (IV or oral) or 3rd generation cephalosporin (IV or IM) on Day 1 (for a maximum of 3 days) followed by randomization to oral ofloxacin for 7 or 14 days. ‡Pyophage and placebo were given in a double-blind manner, while antibiotics were administered open-label.

Supplemental analysis of study characteristics from the literature evaluation: Setting and clinical presentation

Two studies mandated that participants must have fever (both assessed participants with febrile UTI);[9,13] in one study, participants had to have a fever of ≥38°C[9] and in the other, participants were required to have a temperature ≥38°C or <36°C.[13]

Supplemental analysis of study characteristics from the literature evaluation: Study endpoints

Three studies did not specify a TOC timepoint; two evaluated clinical and/or microbiological response as a primary endpoint at Week 6[13] and Day 7,[15] respectively, and one assessed clinical and microbiological response as secondary outcomes at a late response timepoint (7–10 days after EOT)[9]. Among these studies, two did not assess outcomes at later timepoints[9,15] and the third assessed follow-up outcomes at Week 12.[13]

Supplemental analysis of study characteristics from the literature evaluation: Efficacy findings

Among the 13 studies that presented primary endpoints assessing either clinical response, microbiological response, or a composite of both (Table 1), ten studies assessed non-inferiority of the test drug versus comparator on the primary endpoint.[1,7,8,10-13,15,18,20] Eight of these demonstrated non-inferiority: among the seven that assessed composite clinical and microbiological response, the absolute difference between treatment groups varied from –3·4 to 21·2%.[1,8,10-12,18,20] In one of these studies the difference in percentage points varied by the timing of endpoint assessment,[20] and in one, co-primary endpoints of microbiological response were assessed in two different analysis sets, indicating non-inferiority.[12] One of the ten studies concluded that non-inferiority was not met[7] and one concluded that inferiority was met.[13] Three studies that presented primary endpoints assessing clinical and/or microbiological response did not assess either superiority or non-inferiority criteria.[6,17,21]

Table S3: Comparison of clinical and microbiological response for studies that provided the treatment effect for both outcomes at test of cure

Study	Analysis set	Timepoint	Test	Comparator	Clinical response: Treatment difference, percentage points (95% CI)	Microbiological response: Treatment difference, percentage points (95% CI)	Discordant findings (i.e. significant treatment effect for one outcome but not the other)?
Dunne et al. 2023[7]	Micro-mITT	TOC	Sulopenem IV*	Ertapenem IV*	1·0 (–3·1, 5·1)	–6·8 (–12·5, –1·1)	Yes
Eckburg et al. 2022[8]	Micro-ITT	TOC	Tebipenem pivoxil hydrobromide (oral)	Ertapenem IV	–0·6 (–4·0, 2·8)	–4·5 (–10·8, 1·9)	No
Kaye et al. 2022[10]	Micro-mITT	TOC	Cefepime/ enmetazobactam IV	Piperacillin/ tazobactam IV	3·5 (–1·0, 8·0)	19·0 (12·3, 25·4)	Yes
Kaye et al. 2019[11]	Micro-mITT	TOC	ZTI-01 IV	Piperacillin/ tazobactam IV	–0·8 (–7·2, 5·6)	9·6 (–1·0, 20·1)	No
Kaye et al. 2018[12]	Micro-mITT	TOC	Meropenem-vaborbactam IV	Piperacillin-tazobactam IV	4·4 (–2·2, 11·1)	9·0 (–0·9, 18·7)	No
Portsmouth et al. 2018[18]	mITT	TOC	Cefiderocol IV	Imipenem-cilastatin IV	2·39 (–4·66, 9·44)	17·25 (6·92, 27·58)	Yes
Roilides et al. 2023[19]	Micro-mITT	TOC	Ceftolozane tazobactam IV	Meropenem IV	–7·3 (–17·99, 10·05)	–3·0 (–17·13, 17·40)	No
Wagenlehner et al. 2024[1]	Micro-ITT	TOC	Cefepime-taniborbactam IV	Meropenem IV	4·5 (–2·6, 12·6)	11·7 (2·9, 21·0)	Yes
Wagenlehner et al. 2019[20]	Micro-mITT	TOC	Plazomicin IV	Meropenem IV	–1·4 (–7·9, 5·2)	14·9 (7·0, 22·7)	Yes

Data are only shown for studies that provided treatment effects at TOC in the same analysis set. In some cases, data for multiple timepoints were presented in the papers but we focused our assessment on the TOC visits only. Treatment effects with a significant difference based on 95% CI are shaded green. In addition to the studies in the table, two studies didn't specify a TOC visit but provided the following relevant data: Lafaurie et al.[13] showed a significant treatment effect between study arms in both microbiological and clinical response at Week 6 (ITT population). Edlund et al.[9] assessed late clinical response and late microbiological response (7–10 days after finishing antibiotic treatment) and showed no significant treatment effect in clinical response or microbiological response per patient, but a significant treatment effect for microbiological response per urinary pathogen (per protocol population). CI=confidence interval. ITT=intention-to-treat. IV=intravenous. Micro=microbiological. mITT=modified intention-to-treat. TOC=test of cure. *Sulopenem IV followed by oral sulopenem etzadroxil/probenecid or ertapenem IV followed by oral ciprofloxacin or amoxicillin-clavulanate.

Supplemental analysis of patient characteristics from the literature evaluation: Race

Five studies did not report the race of study participants.[6,9,13,15,17] Among the remaining 11 studies, the majority of participants in each study were White (ranging from 76 to 100% of participants across studies).

Supplemental analysis of patient characteristics from the literature evaluation: Risk factors (ORENUC criteria)

*O (**No** risk factors)*
No studies reported inclusion of patients without risk factors. However, one study of children with cUTI reported that 75% of participants did not have a complicating factor,[5] one study of adults with febrile UTI reported that 60% had uncomplicated UTI,[9] and one study of adults with UTI included a subgroup with acute uncomplicated UTI.[17].

*R (**R**ecurrent UTI)*
These risk factors are addressed in the results section in the main paper.

*E (**E**xtraurogenital risk factors)*
By enrolling male patients, all 16 studies included patients with extraurogenital risk factors for cUTI. In addition, nine out of 16 studies reported that patients with diabetes mellitus had been enrolled,[1,6,7,9,10,12,13,17,18] although the severity of the diabetes at enrolment was not provided. Nine studies excluded immunosuppressed and/or immunocompromised patients.[6-8,10-12,17,19,20] One publication stated that the study design allowed immunosuppressed patients to be included (though the proportion of such patients enrolled was not provided);[18] only one study reported the proportion of immunosuppressed patients in the trial population.[13] One study included systemic lupus erythematosus as a possible risk factor for cUTI in the inclusion criteria, although the proportion of patients with systemic lupus erythematosus enrolled in the trial was not provided.[17] No studies reported including patients with neurological disorders such as spinal cord injury.

*N (**N**ephrological risk factors)*
One publication stated that patients with kidney transplant could be included,[18] and nine studies specifically excluded such patients[1,5-8,10-12,20] (one of these also excluded patients with heart and/or lung or pancreatic transplants[1]). In addition, 14 studies excluded patients with kidney failure,[1,5-13,17-20] using exclusion criteria for kidney function values below a certain limit (creatinine clearance ranging from <10 to <60 mL/min or estimated glomerular filtration rate <30 mL/min/1.73m^2) and/or for those receiving dialysis. Two studies did not detail exclusion criteria related to kidney function.[15,21] Ten studies specifically mentioned exclusion of patients with renal/kidney abscess.[1,5-8,10-12,19,20]

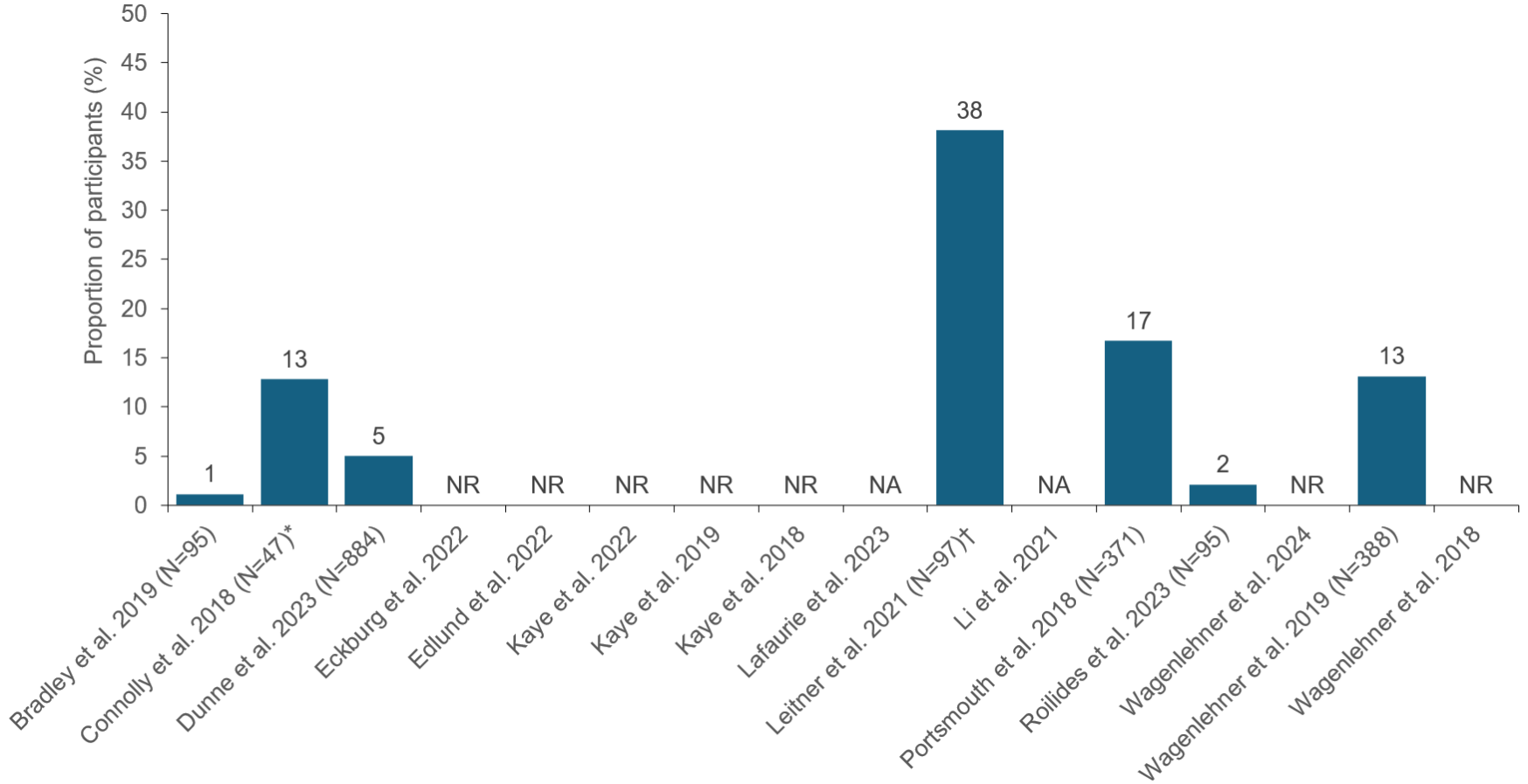
Based on the number of trials that excluded patients with kidney transplant and/or immunosuppressive therapy, five out of 16 studies might have included patients with a kidney transplant.[9,13,15,18,21] These patients are not only immunosuppressed but almost all patients have an impaired anti-reflux mechanism related to the new ostium from the transplanted ureter. Usually, their own malfunctioning kidney and ureter are still in place.

*U (**U**rological risk factors)*
All publications (except for two)[13,21] reported urological risk factors as indicators of cUTI among study inclusion criteria. One study was in men with non-febrile UTI undergoing transurethral resection of the prostate.[15] One of the studies that did not mention urological risk factors in the definition of cUTI was in men with febrile UTI.[13] No studies used a more precise definition than anatomical or functional disorder of the urinary tract.

One study[18] reported the proportion of patients with a medical history of stones in the urinary tract, and one reported the proportion with stones in the urinary tract at baseline[17] (other studies reported the frequency of obstructive uropathy in the study population, or the frequency with removable sources of infection, both of which might include stones[7,10,12,18,19]). No studies provided any information on whether the patients with cUTI had ever been evaluated for UTI by a urologist.

*C (**C**atheters)*
Among the seven studies reporting the proportion of patients with catheters in the study population (Figure S4), one provided the duration of catheterization[15] (two of these studies excluded permanent/chronic indwelling catheters[7,19] and one paper noted the catheters were indwelling within 14 days before screening[18]). No studies reported the proportion of patients enrolled with urinary stents, nephrostomy tubes, prostheses, or foreign bodies in the urogenital tract (but one reported the frequency of patients with recent bladder instrumentation, not further specified).[19] A foreign body in the urinary tract is the most important risk factor for recurrence after cUTI and might significantly skew microbiological sustainability results already from the first week after end of treatment due to persistent biofilm.

Figure S4: Proportion of participants with urinary catheters



Data are presented for all treatment groups combined. Analysis sets used to report data varied between studies. NA=not applicable (patients with catheters were not eligible for study enrolment). NR=not reported (patients with catheters were eligible for inclusion, but the proportion enrolled is not provided). *Data were presented for the subset of patients with cUTI only (not for the overall population, which also include patients with acute pyelonephritis). †Reflects the proportion of patients with an indwelling catheter before transurethral resection of the prostate.

Table S4: Definitions used to assess sustainability of microbiological response

Study	Definition of microbiological response used to assess sustainability of effect	Timepoint used to assess sustainability of effect
Bradley et al. 2019[5]	Microbiological response: All baseline pathogens eradicated	20–36 days after end of treatment (late follow-up visit)
Eckburg et al. 2022[8]	Microbiological response: Reduction in the baseline uropathogen to $<10^3$ CFU/mL and a negative repeated blood culture if the blood culture was positive for a uropathogen at baseline	Day 25 ± 2 days (late follow-up visit)
Kaye et al. 2022[10]	Microbiological eradication: Reduction of qualifying baseline pathogen to $<10^3$ CFU/mL in urine	Day 21 (late follow-up visit)
Kaye et al. 2019[11]	Microbiological eradication: Baseline pathogen reduced to $<10^4$ CFU/mL on urine culture (and negative on repeat blood culture if positive at baseline)	Day 26 ± 2 days (late follow-up visit)
Kaye et al. 2018[12]	Microbiological eradication: Baseline pathogens in urine reduced to $<10^4$ CFU/mL (FDA endpoint) or $<10^3$ CFU/mL (EMA endpoint)	14 days after end of treatment (late follow-up visit)
Portsmouth et al. 2018[18]	Microbiological response: Urine culture $\leq 10^4$ CFU/mL	14 days after end of treatment (follow-up visit)
Wagenlehner et al. 2024[1]	Microbiological success: Reduction of all Gram-negative bacterial pathogens found at baseline to $<10^3$ CFU/mL	Day 28–35 (late follow-up visit)
Wagenlehner et al. 2019[20]	Microbiological eradication: Reduction in the baseline uropathogen from $\geq 10^5$ CFU/mL to $<10^4$ CFU/mL	Day 24–32 (late follow-up visit)
Wagenlehner et al. 2018[21]	Microbiological response: Elimination or reduction of the study entry pathogen(s) to $\leq 10^3$ CFU/mL upon urine culture	Day 24 (end of study visit)

In addition to the studies in this table (which included microbiological response as an endpoint to assess sustainability of effect), one study assessed microbiological recurrence at late follow up (33–47 days after end of treatment) as a criterion for lack of sustainability of treatment effect (defined as a urine culture with $>10^5$ CFU/mL of regrowth of a baseline pathogen that was eradicated at TOC)[6] and one study which evaluated the primary endpoint at Week 6, assessed recurrent UTI between Week 6 and Week 12 (defined as new UTI symptoms, with a positive urine culture [$\geq 10^3$ CFU/mL], and a new antibiotic prescription for UTI).[13] CFU=colony forming units. EMA=European Medicines Agency. FDA=US Food and Drug Administration. TOC=test of cure. UTI=urinary tract infection.

Weighting of risk factors for cUTI identified by the expert working group

Risk factors are presented in order of decreasing risk of causing treatment failure, meaning that urinary catheters cause the highest risk of treatment failure and immunocompromise causes the lowest risk among the factors listed.

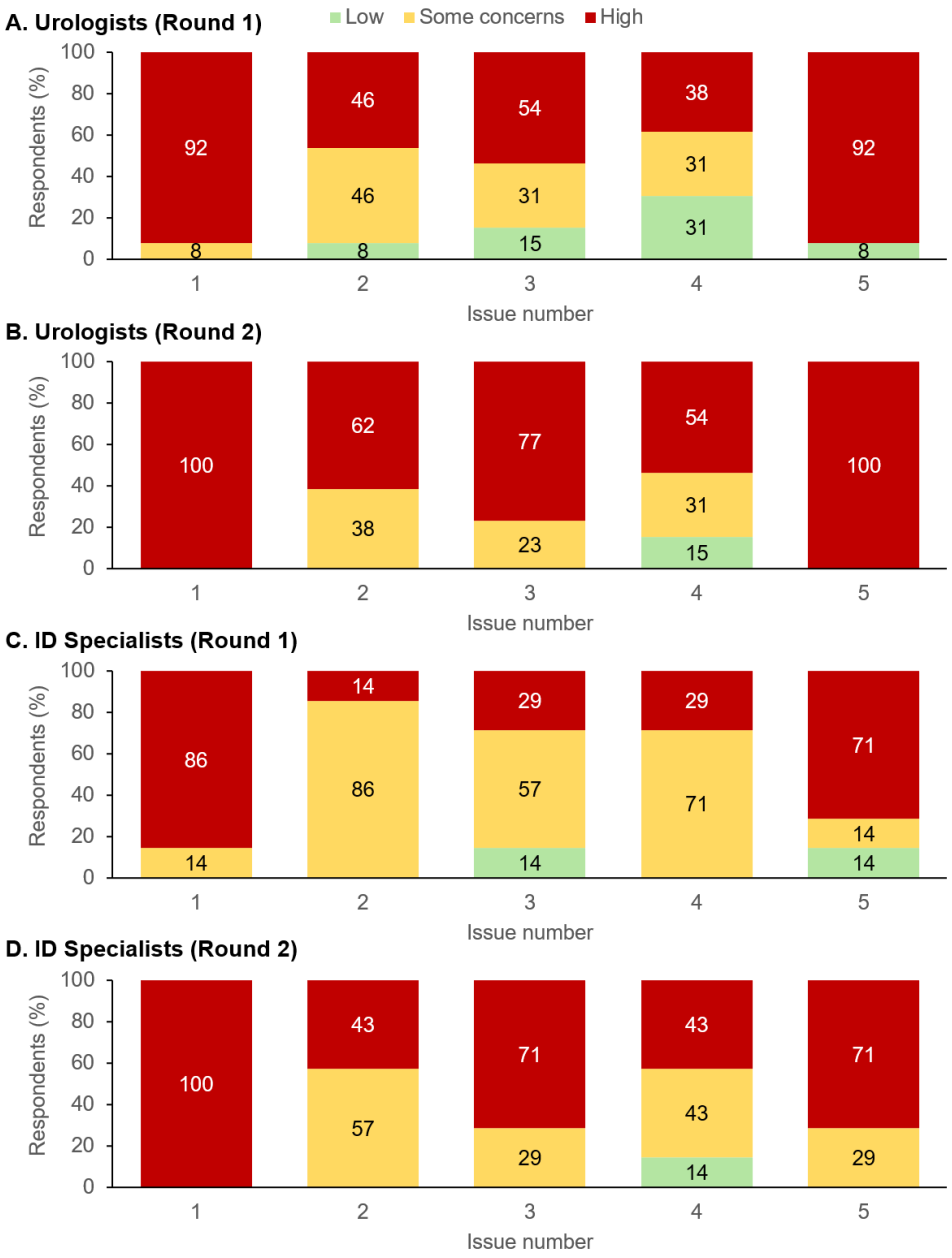
1. Presence of an indwelling urinary/bladder catheters or stents
2. Anatomical abnormalities of the urinary tract causing drainage problems at any level
3. Patients with urinary stones
4. Prior history of symptomatic UTI in the last 6 months
5. Prior history of prostatitis
6. Diabetes mellitus with unstable metabolic situation
7. Pregnant females
8. Antibiotic treatment within previous 30 days
9. Presence of percutaneous nephrostomy
10. Immunocompromised

Table S5: Statistical analysis – Interrelationship of parameters influencing the p-value of falsely rejecting a null hypothesis stating a difference between treatment arms in non-inferiority studies

		Number of patients in control/test groups			
		200/200	**200/150**	**150/150**	**100/100**
Absolute difference in number of patients with treatment success	**Delta = 0**	1%	1.5%	2.3%	5%
	Delta = 5	12%	13%	15%	20%
	Delta = 10	50%	50%	50%	50%

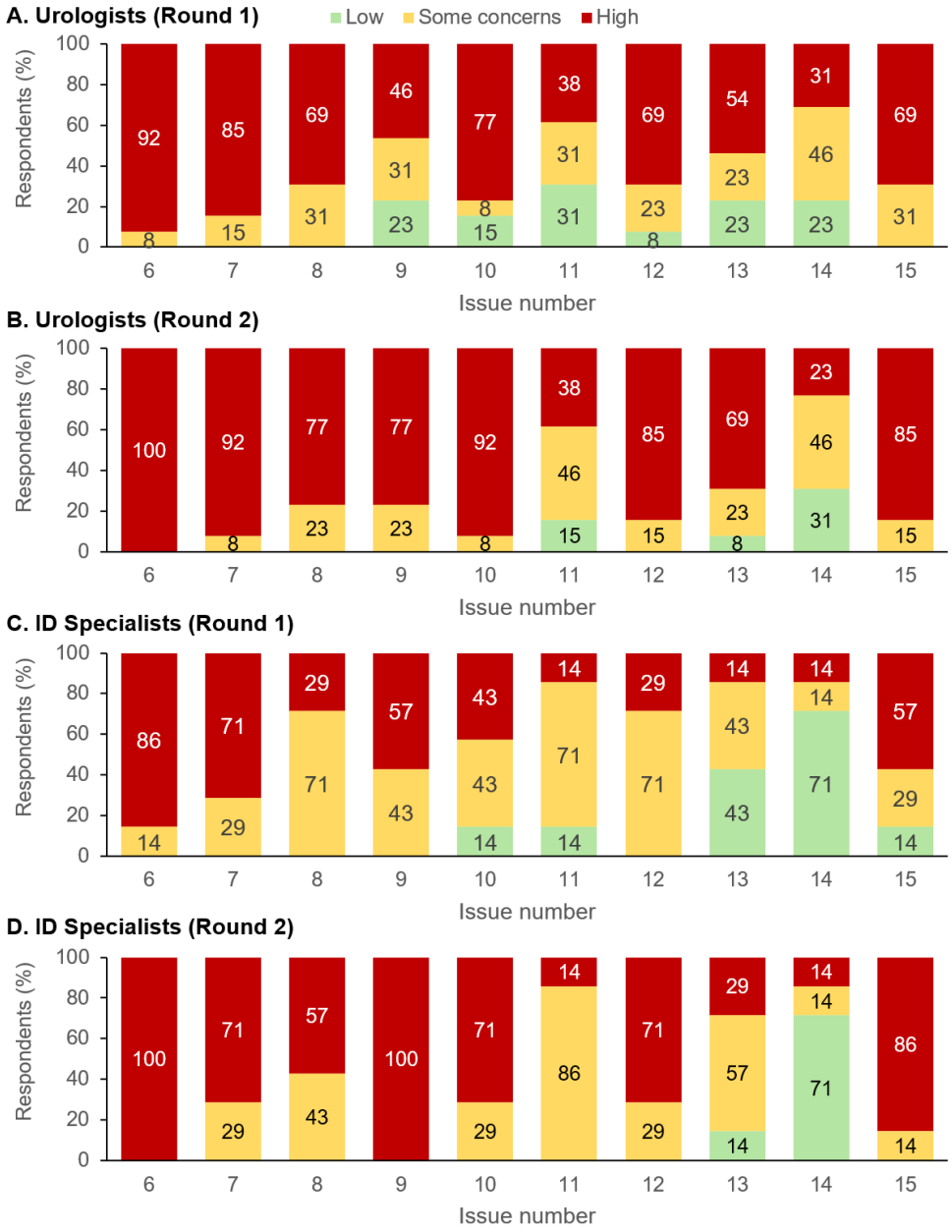
The table shows p-values according to the number of patients in study groups and the difference in the number of patients with treatment success between the two groups. The effect of treatment in the control group is set to 80% and the non-inferiority margin to 0·10. If the non-inferiority criterium is reduced from 10 to 5, the p-value increases from 12% to 20%. If the non-inferiority criterium lies within 5% of the confidence interval for the null hypothesis, the p-value will be less than 5%.

Figure S5: Delphi voting results among urologists (n=13) and ID specialists (n=7) in rounds 1 and 2 for signalling question A: Clinical situation category (Issues 1–5)



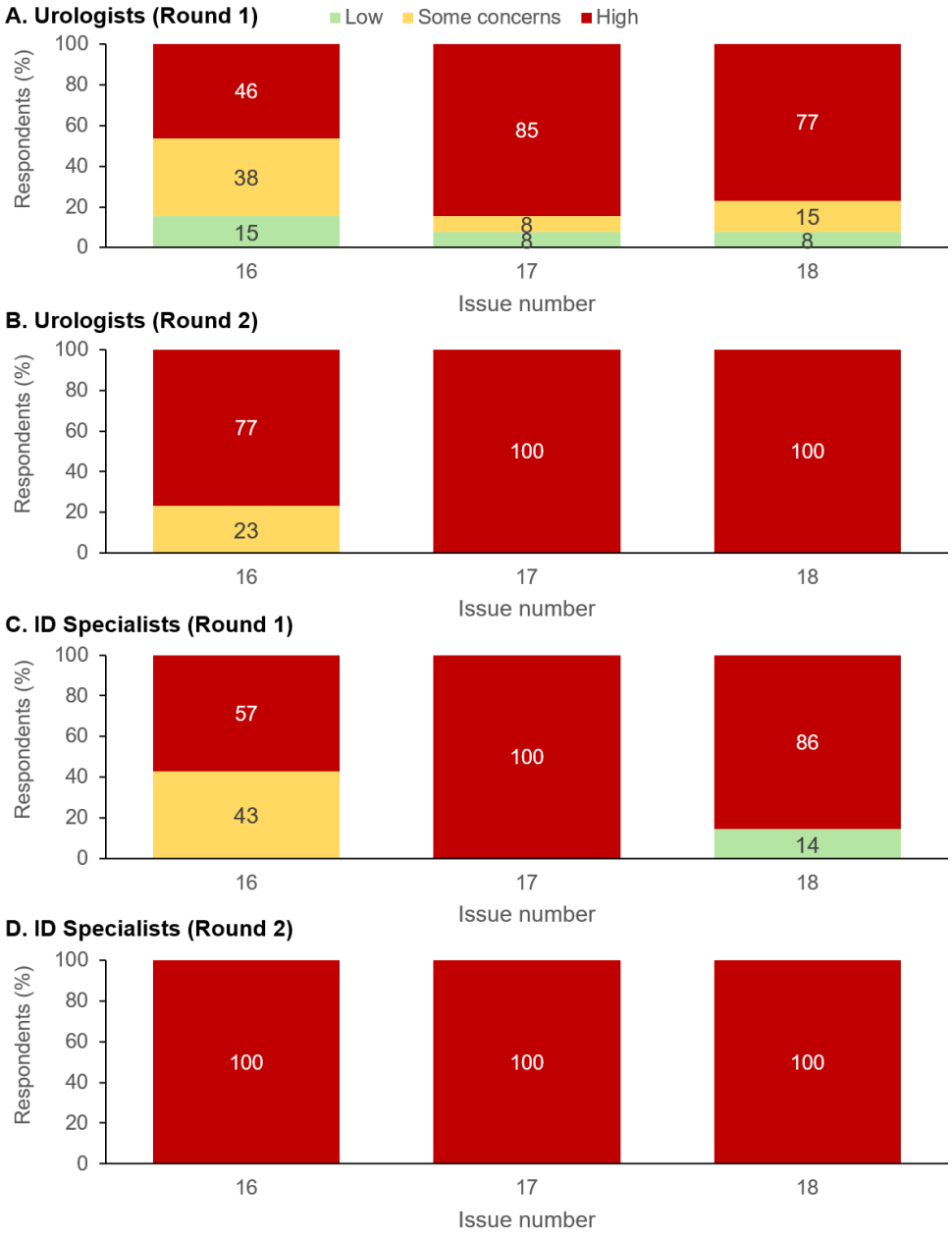
Signalling Question A was "How would you assess the risk of bias in study outcomes in terms of clinical and microbiological success (including sustainability), if these issues (risk factors) were: not considered at all, or not equally distributed, or not clearly defined (i.e., fever) in study groups?". The issues were: 1. The type and frequency of clinical presentation and severity, i.e., percentages of cystitis, non-febrile and febrile pyelonephritis, and urosepsis; 2. Whether the infection was nosocomial- or community-acquired; 3. Fever; 4. The general condition of the patient (ASA status); 5. If two or more of the issues in this category were not satisfactorily described. ASA=American Society of Anaesthesiologists. ID=infectious diseases.

Figure S6: Delphi voting results among urologists (n=13) and ID specialists (n=7) in rounds 1 and 2 for signalling question A: Patient category (Issues 6–15)



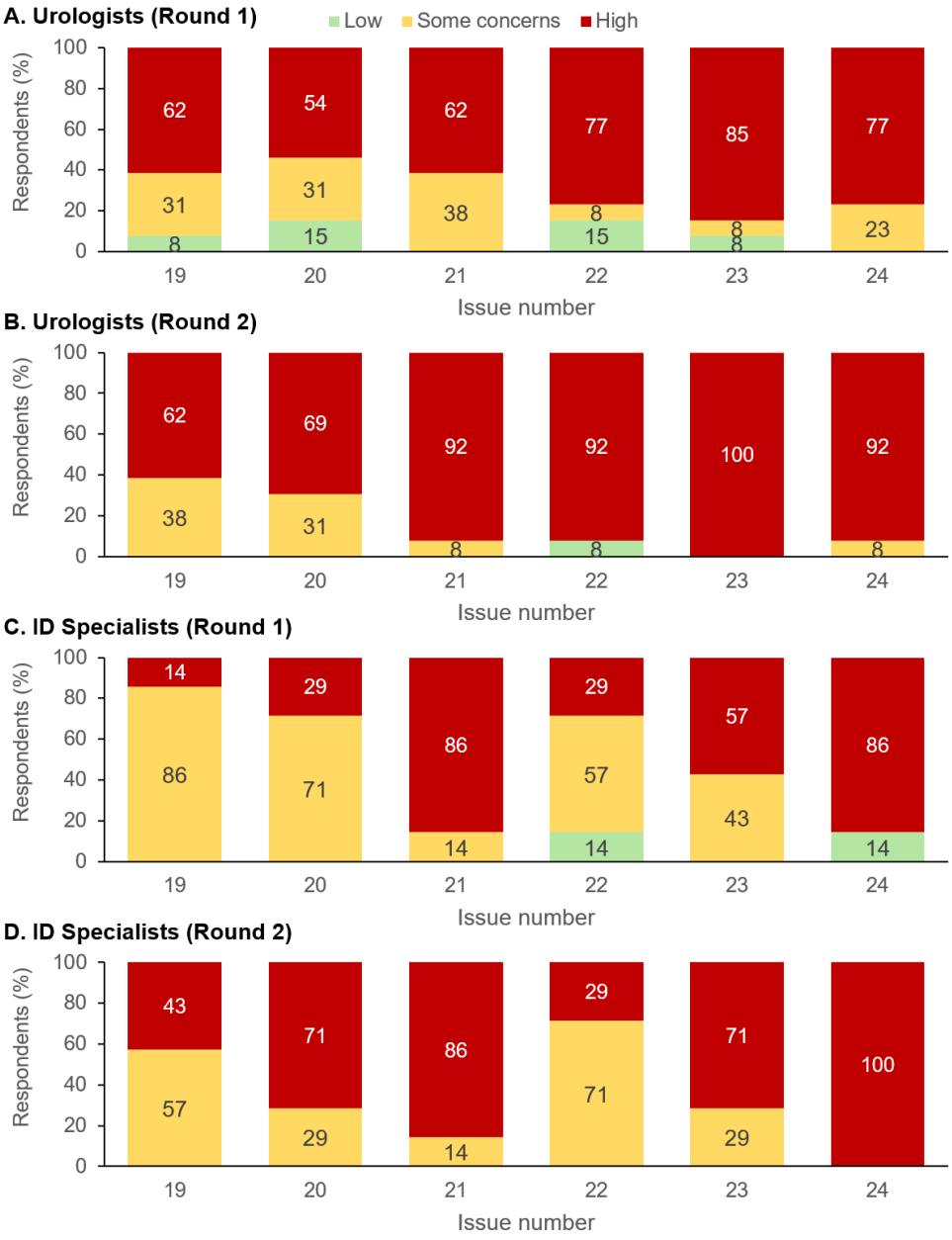
Signalling Question A was "How would you assess the risk of bias in study outcomes in terms of clinical and microbiological success (including sustainability), if these issues (risk factors) were: not considered at all, or not equally distributed, or not clearly defined (i.e., fever) in study groups?". The issues were: 6. Catheter or stent (including nephrostomy tube) in place at diagnosis, during treatment and during follow-up; 7. Presence of urinary stones anywhere in the urinary tract; 8. History of symptomatic UTI in the previous 6 months; 9. A history of bacterial prostatitis; 10. Evidence of immune suppression or diabetes; 11. Female sex; 12. Antibiotic treatment within the previous 30 days; 13. Premenopausal women who have a history of recurrent UTI associated with sexual intercourse; 14. A history of obstipation; 15. If two or more of the issues in this category were not satisfactorily described. ID=infectious diseases. UTI=urinary tract infection.

Figure S7: Delphi voting results among urologists (n=13) and ID specialists (n=7) in rounds 1 and 2 for signalling question B: Pathogen category (Issues 16–18)



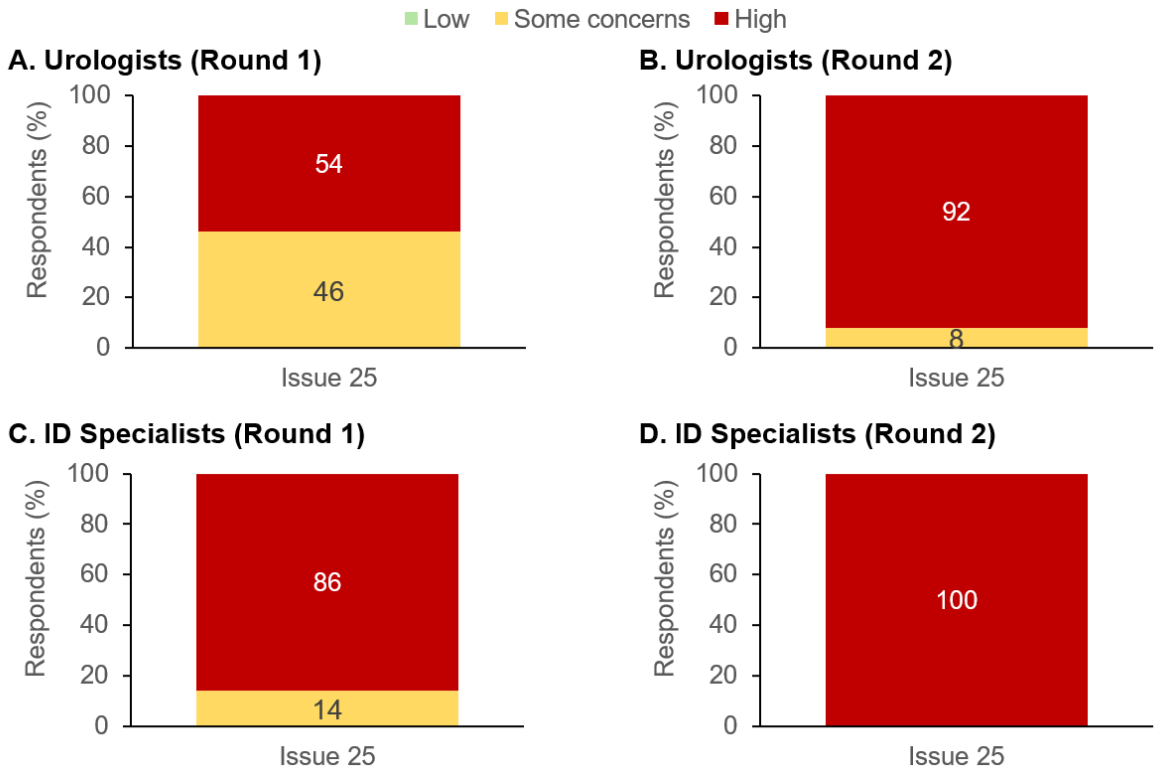
Signalling Question B was "How would you assess the risk of bias of study outcomes in terms of microbiological success, if these issues related to the pathogen were: not reported at all, or not equally distributed in study groups?". The issues were: 16. The spectrum of pathogens; 17. Resistance to study drugs; 18. If both issues in this category were not satisfactorily described. ID=infectious diseases.

Figure S8: Delphi voting results among urologists (n=13) and ID specialists (n=7) in rounds 1 and 2 for signalling question C: Study characteristics category (Issues 19–24)



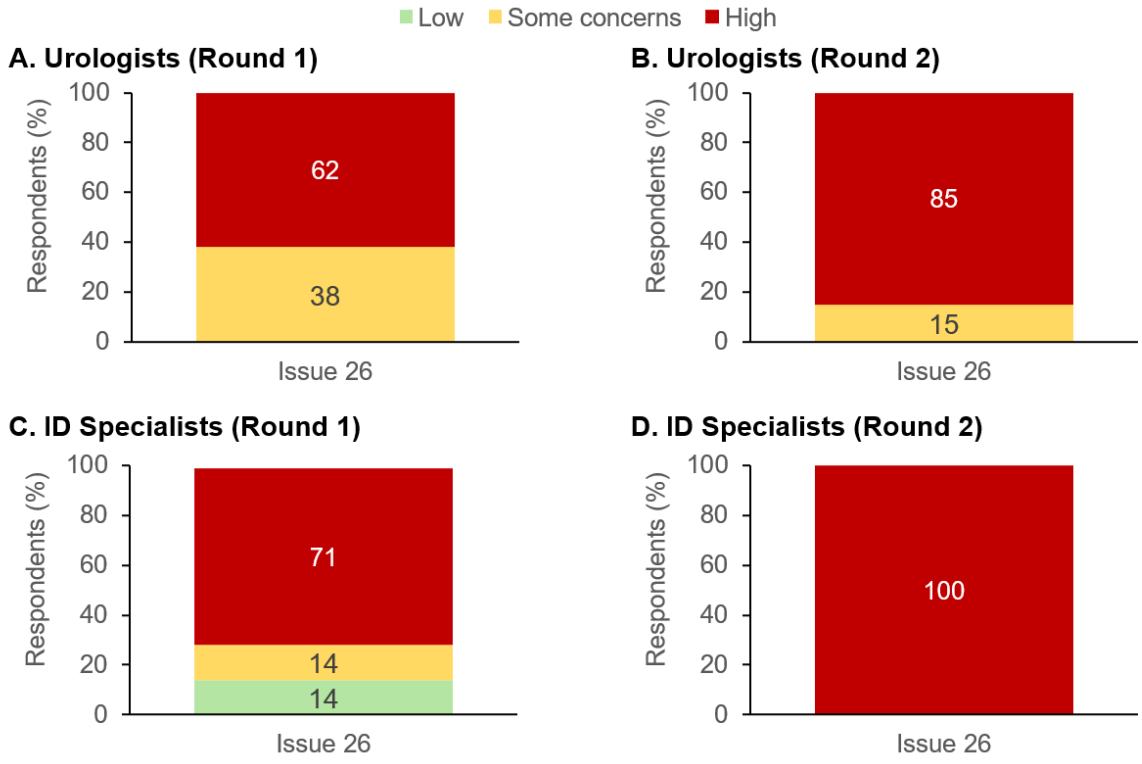
Signalling Question C was "How would you assess the risk of bias when comparing studies if these differences in study characteristics were present?". The issues were: 19. Difference between studies in the duration of treatment with study drug of more than 100% of average treatment duration (typically >3 days); 20. The period from EOT to TOC assessment differs by >1 week between studies; 21. The observation period from EOT to assessment of sustainability of effect differs by >3 weeks between studies; 22. A discrepancy between studies in microbiological success criteria of ≥10 CFU/mL (using $\leq 10^3$ instead of $\leq 10^4$ CFU/mL); 23. A discrepancy between studies in the definition of microbiological "eradication" (using $\leq 10^3$ CFU/mL or no definition at all); 24. If two or more of the issues in this category were present. CFU=colony forming unit. EOT=end of treatment. ID=infectious diseases. TOC=test of cure.

Figure S9: Delphi voting results among urologists (n=13) and ID specialists (n=7) in rounds 1 and 2 for signalling question D: All categories (Issue 25)



Signalling Question D was "How would you assess the overall risk of bias in the outcomes of a study (in terms of clinical and microbiological success)". Issue 25 was: If at least one issue in each of the following categories was present (clinical situation, patient and pathogen)? ID=infectious diseases.

Figure S10: Delphi voting results among urologists (n=13) and ID specialists (n=7) in rounds 1 and 2 for signalling question E: All categories (Issue 26)



Signalling Question E was "How would you assess the overall risk of bias when comparing outcomes between studies (in terms of clinical and microbiological success)". Issue 26 was: If at least one issue in all four categories (clinical situation, patient, pathogen, and study characteristics) was present? ID=infectious diseases.

References

1. Wagenlehner FM, Gasink LB, McGovern PC, et al. Cefepime-Taniborbactam in Complicated Urinary Tract Infection. *N Engl J Med* 2024; **390**: 611-22.
2. Kaul S, Diamond GA. Good enough: a primer on the analysis and interpretation of noninferiority trials. *Ann Intern Med* 2006; **145**: 62-9.
3. Piaggio G, Elbourne DR, Altman DG, Pocock SJ, Evans SJ, Group C. Reporting of noninferiority and equivalence randomized trials: an extension of the CONSORT statement. *JAMA* 2006; **295**: 1152-60.
4. Fleming TR. Current issues in non-inferiority trials. *Stat Med* 2008; **27**: 317-32.
5. Bradley JS, Roilides E, Broadhurst H, et al. Safety and efficacy of ceftazidime-avibactam in the treatment of children ≥3 months to <18 years with complicated urinary tract infection: Results from a phase 2 randomized, controlled trial. *Pediatr Infect Dis J* 2019; **38**: 920-8.
6. Connolly LE, Riddle V, Cebrik D, Armstrong ES, Miller LG. A multicenter, randomized, double-blind, phase 2 study of the efficacy and safety of plazomicin compared with levofloxacin in the treatment of complicated urinary tract infection and acute pyelonephritis. *Antimicrob Agents Chemother* 2018; **62**: e01989-17.
7. Dunne MW, Aronin SI, Das AF, et al. Sulopenem for the Treatment of Complicated Urinary Tract Infections Including Pyelonephritis: A Phase 3, Randomized Trial. *Clin Infect Dis* 2023; **76**: 78-88.
8. Eckburg PB, Muir L, Critchley IA, et al. Oral Tebipenem Pivoxil Hydrobromide in Complicated Urinary Tract Infection. *N Engl J Med* 2022; **386**: 1327-38.
9. Edlund C, Ternhag A, Skoog Ståhlgren G, et al. The clinical and microbiological efficacy of temocillin versus cefotaxime in adults with febrile urinary tract infection, and its effects on the intestinal microbiota: a randomised multicentre clinical trial in Sweden. *Lancet Infect Dis* 2022; **22**: 390-400.
10. Kaye KS, Belley A, Barth P, et al. Effect of Cefepime/Enmetazobactam vs Piperacillin/Tazobactam on Clinical Cure and Microbiological Eradication in Patients With Complicated Urinary Tract Infection or Acute Pyelonephritis: A Randomized Clinical Trial. *JAMA* 2022; **328**: 1304-14.
11. Kaye KS, Rice LB, Dane AL, et al. Fosfomycin for Injection (ZTI-01) Versus Piperacillin-tazobactam for the Treatment of Complicated Urinary Tract Infection Including Acute Pyelonephritis: ZEUS, A Phase 2/3 Randomized Trial. *Clin Infect Dis* 2019; **69**: 2045-56.
12. Kaye KS, Bhowmick T, Metallidis S, et al. Effect of Meropenem-Vaborbactam vs Piperacillin-Tazobactam on Clinical Cure or Improvement and Microbial Eradication in Complicated Urinary Tract Infection: The TANGO I Randomized Clinical Trial. *JAMA* 2018; **319**: 788-99.
13. Lafaurie M, Chevret S, Fontaine JP, et al. Antimicrobial for 7 or 14 Days for Febrile Urinary Tract Infection in Men: A Multicenter Noninferiority Double-Blind, Placebo-Controlled, Randomized Clinical Trial. *Clin Infect Dis* 2023; **76**: 2154-62.
14. ClinicalTrials.gov. NCT02424461: Antibiotic treatment for 7 days versus 14 days in patients with acute male urinary tract infection (PROSTASHORT). 2024. https://clinicaltrials.gov/study/NCT02424461?term=NCT02424461&rank=1 (accessed December 18th 2024).
15. Leitner L, Ujmajuridze A, Chanishvili N, et al. Intravesical bacteriophages for treating urinary tract infections in patients undergoing transurethral resection of the prostate: a randomised, placebo-controlled, double-blind clinical trial. *Lancet Infect Dis* 2021; **21**: 427-36.
16. ClinicalTrials.gov. NCT03140085: Bacteriophages for treating urinary tract infections in patients undergoing transurethral resection of the prostate. 2024. https://clinicaltrials.gov/study/NCT03140085?term=NCT03140085&rank=1 (accessed December 18th 2024).
17. Li Y, Yin Y, Peng X, et al. A randomized, active-controlled, multicentre clinical trial to evaluate the efficacy and safety of oral sitafloxacin versus levofloxacin in Chinese adults with acute uncomplicated or complicated urinary tract infection. *Ann Med* 2021; **53**: 217-26.
18. Portsmouth S, van Veenhuyzen D, Echols R, et al. Cefiderocol versus imipenem-cilastatin for the treatment of complicated urinary tract infections caused by Gram-negative uropathogens: a phase 2, randomised, double-blind, non-inferiority trial. *Lancet Infect Dis* 2018; **18**: 1319-28.
19. Roilides E, Ashouri N, Bradley JS, et al. Safety and Efficacy of Ceftolozane/Tazobactam Versus Meropenem in Neonates and Children With Complicated Urinary Tract Infection, Including Pyelonephritis: A Phase 2, Randomized Clinical Trial. *Pediatr Infect Dis J* 2023; **42**: 292-8.
20. Wagenlehner FME, Cloutier DJ, Komirenko AS, et al. Once-Daily Plazomicin for Complicated Urinary Tract Infections. *N Engl J Med* 2019; **380**: 729-40.
21. Wagenlehner F, Nowicki M, Bentley C, et al. Explorative Randomized Phase II Clinical Study of the Efficacy and Safety of Finafloxacin versus Ciprofloxacin for Treatment of Complicated Urinary Tract Infections. *Antimicrob Agents Chemother* 2018; **62**.