Clinical Microbiology and Infection 31 (2025) 1972–1979

Contents lists available at ScienceDirect

Clinical Microbiology and Infection

journal homepage: www.clinicalmicrobiologyandinfection.org

Systematic review

Nomenclature, definitions, and methodological approaches to estimate the association between antimicrobial treatment and clinical outcomes of drug-resistant bloodstream infections

Nasreen Hassoun-Kheir [1,*], Miriam Roncal Redin [1,2], Appiah-Korang Labi [3], Michael Loftus [4], Andrew J. Stewardson [5], Stephan Harbarth [1], Leonard Leibovici [6], Mical Paul [7], Koen B. Pouwels [8], Mike Sharland [9], Alexander M. Aiken [10], Marlieke E.A. de Kraker [1]

1) Infection Control Program, Geneva University Hospitals and Faculty of Medicine, World Health Organization Collaborating Center, Geneva, Switzerland
2) Department of Preventive Medicine and Public Health, Ramón y Cajal Hospital and Ramón y Cajal Health Research Institute (IRYCYS), Madrid, Spain
3) Department of Medical Microbiology, University of Ghana Medical School, Accra, Ghana
4) Department of Infectious Diseases, Monash University, Melbourne, Victoria, Australia
5) Department of Infectious Diseases, The Alfred and School of Translational Medicine, Monash University, Melbourne, Victoria, Australia
6) Department of internal medicine E, Rabin Medical Center, Tel Aviv, Israel
7) Infectious Diseases Unit, Rambam Healthcare Campus, Haifa, Israel
8) Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford, UK
9) Paediatric Infectious Diseases Service, St George's Healthcare National Health Service Trust, London, UK
10) Department of Infectious Disease Epidemiology and International Health, London School of Hygiene and Tropical Medicine, London, UK

A R T I C L E I N F O

Article history:
Received 20 May 2025
Received in revised form
22 July 2025
Accepted 31 July 2025
Available online 6 August 2025

Editor: A. Kalil

Keywords:
Analytical methodology
Antimicrobial resistance
Bias
Bloodstream infection
Causal inference
Inappropriate initial antibiotic treatment

A B S T R A C T

Background: Antimicrobial resistance increases the risk of misaligned initial antibiotic treatment (IAT), as susceptibility data are typically delayed. The causal effect on patient outcomes, however, is unclear due to reliance on observational studies with methodological heterogeneity.
Objectives: To describe the terminology and definitions for IAT misalignment and evaluate methods used to analyse its association with mortality and hospital length of stay (LOS) for patients with drug-resistant bloodstream infections (BSIs).
Methods: A systematic review.
Data sources: PubMed and EMBASE: January 1990 to August 2024.
Study eligibility criteria: We included studies on drug-resistant BSIs caused by ESKAPE pathogens (*Enterococcus faecium, Staphylococcus aureus, Klebsiella pneumoniae, Acinetobacter baumannii, Pseudomonas aeruginosa, Enterobacter* species, and other Enterobacterales). Eligible studies defined IAT misalignment and assessed its effect on mortality/LOS.
Participants: Patients with drug-resistant BSIs.
Exposure: (mis)aligned IAT.
Assessment of risk of bias: Revised versions of the Joanna Briggs Institute tools.
Methods of data synthesis: Qualitative synthesis.
Results: From 3627 screened publications, 187 studies were included, predominantly cohort studies ($n = 183$). The most common terminology for IAT misalignment was “(in)appropriate” ($n = 139$, 74.3%), followed by “(in)adequate” ($n = 34$, 18.2%). Definitions primarily considered *in vitro* susceptibility to prescribed antibiotic(s) ($n = 184$, 98.4%), with up to nine additional criteria. Impact of (in)appropriate IAT on mortality ($n = 186$) was mostly evaluated using logistic or Cox regression, including various confounder selection methods, showing an association in 122 of 186 studies (65.6%). Admission-to-infection time and infection-to-treatment time were rarely considered. Impact of (in)appropriate IAT on LOS was shown in two of nine studies. Only four studies explicitly analysed postinfection LOS. No study scored a low risk of bias, due to residual confounding and time-dependent bias.
Discussion: Wide variability of IAT definitions and impact analysis was observed, with a high risk of bias, hindering data aggregation and limiting understanding of the causal effect of inappropriate IAT on

* Corresponding author: Nasreen Hassoun-Kheir.
E-mail address: nasreen.hassounkheir@gmail.com (N. Hassoun-Kheir).

https://doi.org/10.1016/j.cmi.2025.07.033

clinical outcomes. Guidelines are required to improve study quality and harmonize future research.
Nasreen Hassoun-Kheir, Clin Microbiol Infect 2025;31:1972

Background

The burden of infectious diseases is aggravated by the emergence of antimicrobial resistance (AMR) [1,2]. AMR may increase the risk of ineffective initial antibiotic treatment (IAT), as IAT is mostly prescribed before the results of microbiological testing are available. Consequently, treatment of an individual patient might not align with the identification and/or susceptibility profile of the infecting pathogen. This misalignment is described using different names, like "inappropriate", "inadequate", or "ineffective" IAT. As it can usually only be determined retrospectively, assessment of its association with patients' outcomes is mostly evaluated in observational studies. Although an increased risk of ineffective IAT is generally assumed to be one of the key mechanisms by which AMR impacts patient outcomes, estimating the causal effects of IAT on patients' outcomes using observational data is challenging.

In a recent systematic review examining the impact of inappropriate IAT for severe bacterial infections, some studies reported an association between misaligned treatment and excess mortality or length of hospital stay, whereas others did not find such a relationship [3]. The authors describe this discrepancy but do not provide methodological details on the included studies to explore the origin of these contractionary findings. It might be partially explained by differences in the definitions of IAT misalignment or by variation in analytical approaches and methods used across the different studies [4]. For example, immortal time bias—arising when patients are considered treated before true treatment initiation—can result in misleading findings. Other challenges in this setting include confounding by indication and collider stratification bias (Table 1) [5–8], which can challenge the conditional exchangeability (or 'no unmeasured confounding') assumption, which is required for causal effect estimation. Conditional exchangeability refers to a situation where, after correctly adjusting for all confounding factors, the treatment groups are comparable—as if individuals had been randomly assigned to treatment or control [9]. Consistency is also an important requirement for causal inference. It emphasizes the need for an exposure to be clearly defined, which, in this case, would include, for example, the selection of antibiotics considered, to ensure that the exposure has the same effect in an observational study as it would have had in a trial [10–12]. Positivity is also a necessary requirement, which, in this context, refers to the probability of receiving (mis)aligned IAT, which should be nonzero for every included patient. Finally, there should be no interference (exposure of one individual should not affect the potential outcome of another) [13]. These crucial elements are often overlooked in studies that aim to estimate the causal effect of (mis)aligned IAT on clinical outcomes.

In this systematic review, we catalogued the terminology used to describe misalignment of IAT in patients with bloodstream infections (BSIs), assessed the definitions used for IAT misalignment, and described the analytical methods used to determine its association with clinical outcomes.

Methods

Eligibility criteria

A review protocol was registered in the International prospective register of systematic reviews (PROSPERO) (CRD42022286778) [20]. Eligibility criteria for study selection were defined using the PICOS (Population, Intervention, Comparison, Outcomes, and Study design) framework [21]. Patient population was hospitalized patients with BSIs caused by ESKAPE pathogens: *Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa*, *Enterobacter* spp., or other Enterobacterales. Exposure of interest was IAT, comparing aligned with misaligned IAT, with IAT defined as treatment started within 24–48 hours of BSI onset, which could include empirical as well as targeted treatment. Studies should report a definition for the alignment of IAT given to treat BSI. Outcomes included mortality and/or hospital length of stay (LOS), and studies had to clearly describe the methods used to determine the association between IAT and these clinical outcomes. Eligible study designs were observational studies (cohort, case-control, and cross-sectional) and interventional studies (randomized and nonrandomized). Systematic reviews were used for the purpose of study identification. As we were interested in methods for causal estimation of the association between IAT and patients' outcomes, we excluded studies that only

Table 1
Common types of bias and confounding in studies assessing the impact of initial antibiotic treatment on patients' outcomes

Bias type	Description	Direction of bias in studies on IAT and clinical outcomes	Recommended reading
Confounding by indication	Occurs when the reason (e.g. severity of illness) for prescribing a certain treatment is also associated with the outcome.	Using broad spectrum antibiotics, with a higher probability of effectiveness, specifically for more severely ill patients would underestimate the effect of IAT.	[14,15]
Collider stratification bias	Happens when conditioning on a variable (collider) influenced by both treatment and outcome (or an ancestor of outcome).	Measuring severity of illness after infection, while analysing it as a baseline variable, would underestimate the association between IAT and clinical outcome.	[16,17]
Immortal time bias	Arises when a period of time after study enrolment is misclassified as "exposed" or ignored in the analysis.	Typically, it underestimates the effectiveness of treatment by only including patients who survive long enough to receive aligned IAT.	[18,19]

IAT, initial antibiotic treatment.

provided univariable analyses of risk factors/predictors for a clinical outcome of interest without further analyses.

Search strategy

The EMBASE and Medline (PubMed) databases were searched on August 23, 2024, using a search strategy including the following concepts: (ESKAPE pathogens AND resistance AND bloodstream infection AND outcome (mortality OR length of stay [LOS]) AND different terms of inappropriate antibiotic prescribing (Table S1). We searched publications from 1990 onwards, with no language restriction. A systematic reference search of previous important publications in the field was performed [3,4,22].

Study selection process

The total list of publications was extracted and deduplicated using Covidence software [23]. A screening tool was developed for each stage of screening and piloted among all reviewers. Single reviewing of titles and abstracts and then of selected full texts was completed by four independent reviewers (N.H.K., M.L., A-.K.L., and M.R.R.); any uncertainties were resolved by discussion during weekly meetings that were held among the four reviewers. The selection process was summarized in a preferred reporting items for systematic reviews and meta-analyses flow diagram (Fig. 1).

Data collection

Single data extraction was completed by a team of seven reviewers (N.H.K., M.L., A-.K.L., M. E.A.d.K., A.A., M.R.R., and A.J.S.). Extracted data elements included study design, study setting, population characteristics, causative pathogen(s), resistance trait (s), definitions of BSI onset, terminology used to describe alignment of IAT, elements used for IAT definition, outcome measures (mortality and/or hospital LOS), and analytical methods used to determine the impact of misaligned IAT on clinical outcomes. Meetings were held to discuss difficulties in data extraction to ensure harmonized data extraction.

Study risk of bias assessment

Single risk of bias assessment was performed by seven reviewers (N.H.K., M.L., A-.K.L., M.E.A.d.K., A.A., M.R.R., and A.J.S.), using the Joanna Briggs Institute critical appraisal tools for included study designs [24]. Study scores were calculated as the percentage of positive responses out of the maximal score, and studies were separated into low (≥75%), medium (≥50–74%), and high risk of bias (<50%). As the specific Joanna Briggs Institute domains on bias identification and strategies for addressing confounding were crucial for the study objectives, these were further specified into sub-domains, which all needed to be fulfilled in a study to be classified as a low risk of bias study (Table S2).

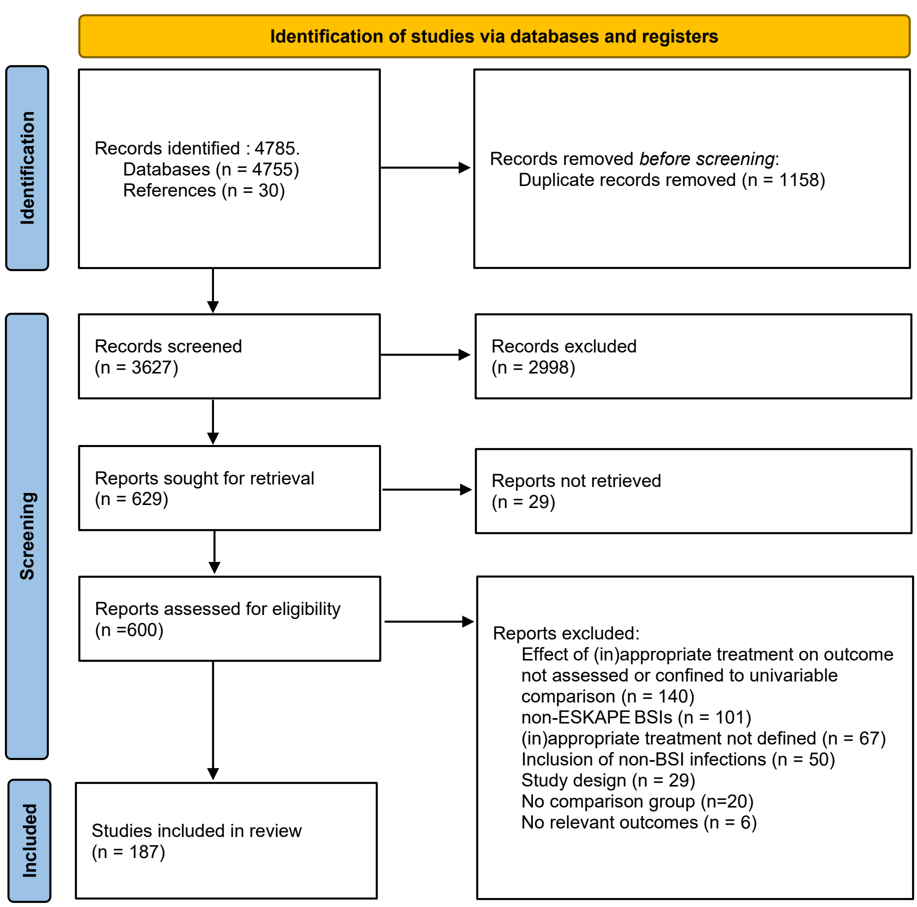

Fig. 1. The preferred reporting items for systematic reviews and meta-analyses flowchart for the systematic review. BSIs, bloodstream infections; ESKAPE, *Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa*, *Enterobacter* species, and other Enterobacterales.

Synthesis methods

Study characteristics were summarized descriptively without meta-analysis. A critical review of the nomenclature, definitions, and analytical methods used to analyse the effectiveness of treatment in the included studies was performed.

Results

Study selection

A total of 3627 nonduplicate publications were screened by title/abstract. Of these, 629 full texts were assessed, and 187 studies fulfilled the inclusion criteria (Fig. 1; Table S3). The main reasons for exclusion were studies not analysing the associations between IAT and a relevant clinical outcome ($n = 140$) and studies reporting on BSIs due to non-ESKAPE pathogens ($n = 101$).

Study characteristics

All included studies were observational studies ($n = 187$); 183 cohort studies (144 retrospective, 34 prospective, and five undetermined), three case-control studies, and one cross-sectional study, published during the period 1995–2024. Most were single-centre studies ($n = 117$, 62.6%) and included all hospital departments ($n = 137$, 73.3%). Thirteen studies (7.0%) focused on intensive care units (11 adult and two neonatal intensive care units), and eight studies exclusively included patients with haematological diseases (4.3%).

Across all studies, a total of 135 699 patients were included, with a median sample size of 211 patients (interquartile range (IQR): 114–377) per study. Most studies included BSIs due to both susceptible and multidrug-resistant pathogens ($n = 134$, 71.7%), the remainder of the studies only included multidrug-resistant pathogens. Sixty-six percent of the included studies ($n = 124$ studies) focused on BSI due to a single ESKAPE pathogen, with *S. aureus* ($n = 34$) and *A. baumannii* ($n = 33$) as the leading pathogens. Although most studies included both community- and hospital-acquired infections ($n = 144$, 77%), 24 studies (12.8%) focused exclusively on hospital-acquired BSIs, and 19 studies (10.2%) examined only community-acquired/healthcare-associated community-onset BSIs.

Nomenclature and definitions of misaligned initial antibiotic treatment

The association between treatment misalignment and clinical outcomes was mentioned in the title of 52 of 187 study publications (27.8%). The most frequently used nomenclature to describe the phenomenon of "misaligned" individual antibiotic treatment was "(in)appropriate" treatment ($n = 139/187$, 74.3%), followed by "(in)adequate" treatment ($n = 34$, 18.2%). "(In)active" was used in five studies (2.7%), whereas "(dis)concordant", and "(in)effective" terms were used in four studies each (2.1%). Last, "(in)correct" treatment was used in one study. Hence, the term (in)appropriateness is used in the remainder of this manuscript. Most studies provided a definition of what was considered appropriate treatment ($n = 147$, 78.6%), 16 studies specified only what was considered inappropriate (8.6%), and 24 studies provided both definitions (12.8%). The elements used to define the appropriateness of antibiotic treatment varied widely (Fig. 2). *In vitro* susceptibility of the infecting bacteria to the antibiotic agent was used in 184 of 187 studies (98.4%). A prespecified time window for initiating IAT to be considered appropriate was defined in 92 studies ($n = 92$, 49.2%), mostly ranging between 24 and 48 hours after BSI diagnosis. Few studies specified a minimum duration of such a treatment to be considered appropriate ($n = 26$, 13.9%). Few studies evaluated pharmacokinetic/pharmacodynamic criteria ($n = 12$, 6.4%). Across 112 studies that included polymicrobial infections, only 14 studies required appropriate IAT to include coverage of all isolated pathogens (12.5%).

BSI onset was defined as the time of blood culture collection in 89 studies (47.6%), the time of reporting a positive culture result in 14 studies (7.5%), and the onset of signs/symptoms in three studies (1.6%). Notably, 80 studies (42.8%) did not provide a clear definition for BSI onset. The "initial treatment phase" was defined as treatment given within a specified time window (mostly ranging between 24 and 72 hours), regardless of the availability of antibiotic susceptibility testing results in 77 studies (41.1%), whereas 56 studies (30.0%) considered the "empiric treatment phase", which included treatment provided until antibiotic susceptibility testing results were available. The remaining studies used a combination of these two approaches.

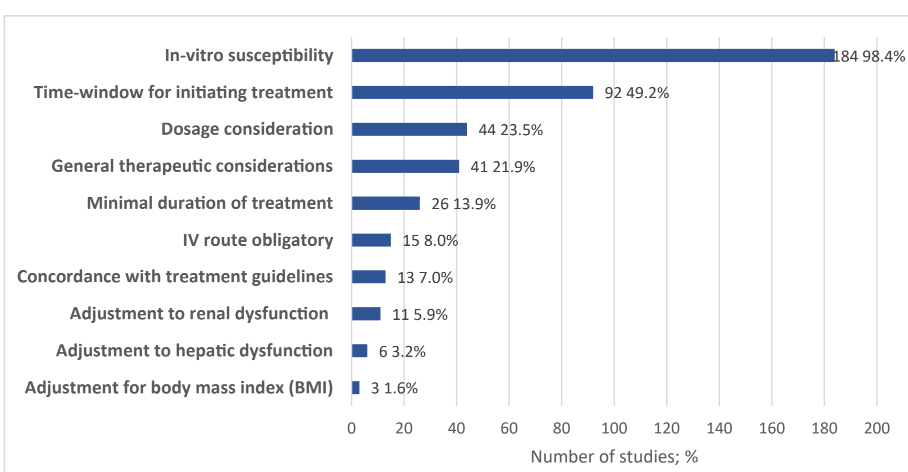

Fig. 2. Elements included in the definition of appropriateness of antibiotic treatment in scientific literature assessing the impact on clinical outcomes in patients with drug-resistant ESKAPE (*Enterococcus faecium, Staphylococcus aureus, Klebsiella pneumoniae, Acinetobacter baumannii, Pseudomonas aeruginosa, Enterobacter* species, and other Enterobacterales) bloodstream infections ($N = 187$).

Analysing the impact on mortality

Mortality was assessed in 186 studies at various time points (Fig S1). The most common time point for mortality assessment was 28–30 days after BSI diagnosis (n = 120, 64.2%), for which most studies did not include post-discharge follow-up (102/120, 85.0%). It was pursued in 20 studies (18 with 28- to 30-day mortality, one with 14-day, and one with 7-day mortality outcomes). Of these, 14 studies specified the exact follow-up method (e.g. use of death registries, follow-up clinic visits, and telephone calls). The second most common time point for mortality assessment was during hospital stay (n = 50, 26.9%).

Multivariable regression was the most common method used to analyse the association between treatment (in)appropriateness and mortality (n = 182 studies, 97.3%). Among these studies, logistic regression was most frequently chosen (n = 142 studies, 78.0%). In 75 of 142 (52.8%) studies using logistic regression models, post-discharge follow-up was not performed for 30-day mortality. Cox regression was the second most common method used (n = 42 studies, 23.1%), including two studies that also performed logistic regression. Among the 182 studies using multivariable regression, propensity scores (PS) were used in 13 studies (7.0%), either through matching, inverse probability weighting, or including PS as a covariate in a regression model. Several variable selection strategies were used (Table 2). Finally, treatment (in)appropriateness was included in the multivariable model for mortality in 138 of 182 studies; mostly selected based on a predefined selection criterion (n = 99 studies, p value in univariable analysis/another criterion) or forced into the final model (n = 39 studies).

Table 2
Features of multivariable regression models in studies assessing the impact of initial antibiotic treatment (in)appropriateness on mortality in patients with bloodstream infections (n = 182)

Feature[a]	No. of studies (%)
Variable selection method	
Based on p value	140 (76.9)
Clinical reasoning	42 (23.1)
Change in effect estimate	4 (2.2)
Directed Acyclic Graphs	1 (0.5)
Variable inclusion strategy	
Backward selection	41 (22.5)
Stepwise selection/unclear direction	22 (12.1)
Forward selection	19 (10.4)
Best subset method	3 (1.6)
No selection	11 (6.0)
Not reported	90 (49.5)

[a] Several studies applied more than one variable selection method/inclusion strategy.

Table 3
Confounders considered in multivariable analysis assessing the impact of initial antibiotic treatment (in)appropriateness on mortality in patients with bloodstream infections (n = 182)

Confounding domain	No. studies[a] (%)
Comorbidities	117 (64.3)
Demographics	75 (41.2)
Hospital setting (ICU/medical/surgical, etc.)	35 (19.2)
Severity of illness	
Measured before infection onset	44 (24.2)
Unclear timing for measurement	49 (26.9)
Measured after infection onset[b]	60 (32.3)

ICU, intensive care units.
[a] Out of 182 studies applying multivariable regression models.
[b] Including on the day of diagnosis.

Regarding adjustment for confounders, most studies assessing mortality considered patients' comorbidities and demographics. Adjusting for severity of illness was also common, although it was frequently assessed at an unknown time point or after infection onset while including it as a fixed baseline variable in multivariable models (Table 3).

Among studies that exclusively included patients with hospital-acquired BSIs (n = 24), admission-to-infection time interval was addressed only in one study by including time as a fixed covariate in the regression model [25], and none of these studies acknowledged time from infection to appropriate treatment. When considering all studies with some hospital-acquired BSI patients or without reporting on onset of infection (n = 168 studies), admission-to-infection time interval was addressed in 13 studies (7.4%), and was most frequently included as a fixed covariate in multivariable regression analysis (n = 6 studies), or addressed as a time-varying variable in two studies [26,27]. Other time-varying variables were considered in the analysis of five studies: treatment effectiveness in four studies [28–31], and severity of illness in one study [26]. Similarly, infection onset to appropriate treatment time interval was addressed in nine studies, either by including appropriate treatment as a time-varying variable [28–30] or time to appropriate treatment as a fixed variable in the multivariable model [31–36]. Finally, treatment (in)appropriateness was significantly associated with mortality in 122 studies (65.6%)—only in crude analysis (n = 19), only in adjusted analysis (n = 15), or both forms of analyses (n = 88).

Analysing the impact on hospital LOS

The association between treatment (in)appropriateness and hospital LOS was evaluated in nine studies. LOS was defined as the time after infection onset (n = 4), or total LOS including time before BSI (n = 3), and two studies did not provide a clear definition. The analytical methods used were univariable t test/Wilcoxon rank-sum (n = 3), or multivariable techniques; linear regression (n = 3), Cox regression (n = 1), inverse probability weighted comparison (n = 1) [37], and “multivariable methods” were reported in one study. Three multivariable models included treatment (in)appropriateness in the final model [27,38,39]. Admission-to-infection onset time interval was addressed in two studies by including infection as a time-varying variable or as an interval in multivariable regression analysis [26,27]. None of the studies included time-varying confounders or addressed time from infection onset to the initiation of appropriate antibiotic treatment in the analysis. Finally, treatment (in)appropriateness was found to have an association with LOS in two of nine studies (by crude analysis [40] and by both crude and adjusted analyses [38]).

Risk of bias assessment

According to the crude risk of bias assessment scores, 53.5%, 40.1% and 6.4% of the studies scored low (n = 100), medium (n = 75), and high (n = 12) risk of bias, respectively. None of the studies, however, scored low risk of bias on all the a priori defined critical items dealing with identification of confounders, time-dependent bias, and methods for confounding adjustment. Thus, finally, no studies could be considered low risk of bias.

A particularly problematic aspect of these studies was the handling of time in the hospital before BSI for hospital-acquired infections. This was only acknowledged in the analysis of 13 of 187 (7.0%) studies (of which 168 studies may have included hospital-acquired infections). Time between infection onset and appropriate treatment was only acknowledged in 17 of 187 (9.1%) studies. Baseline confounders often included covariates that were measured after the exposure of interest (infection onset), without applying the correct analytical methodology, resulting in a high risk of bias in the domain of time-fixed confounders. Only 56 of 187 (29.9%) studies included the predefined minimal set of confounders in multivariable analysis. Conversely, a low risk of information bias was encountered in included studies—175 of 187 (93.6%) scored low risk of bias in exposure measurement, and 159 of 187 (85.0%) scored low risk of bias in outcome classification.

Discussion

We identified 187 studies that defined misalignment of IAT for BSI due to ESKAPE pathogens and assessed the impact of such treatment on clinical outcomes. The term (in)appropriate treatment was predominantly used with "appropriate" typically defined in the positive sense. BSI onset was frequently considered as the time of blood culture collection ($n = 89$ studies, 47.6%), but was otherwise not, or poorly, defined. The terms "initial" and "empiric" treatments were frequently used interchangeably across studies, despite not being strictly synonymous. Although empiric treatment is, by definition, based on clinical judgement and probabilistic reasoning prior to microbiological confirmation, initial treatment refers to the first antibiotics given within a specific time window. The main criterion for appropriate IAT was *in vitro* susceptibility, and other relevant factors like dosing, timing, pharmacokinetics/pharmacodynamics, and clinical considerations were often overlooked. Some studies included alignment with treatment guidelines as a criterion to retrospectively assess appropriateness, thereby incorporating antibiotic stewardship considerations. Multivariable analysis was frequently applied to determine the causal impact of inappropriate IAT, but in many studies, residual confounding and time-dependent bias could not be ruled out.

Previous reviews on the impact of appropriate antibiotic treatment on patient outcomes focused on pooling effect estimates. One review estimated a 56% reduction in odds of mortality (OR: 0.44, 95% CI: 0.37–0.52) associated with appropriate treatment in a subgroup of patients with BSIs and sepsis based on 63 studies [3]. In another recent systematic review ($n = 200$ studies) focusing on BSIs, inappropriate empirical treatment (defined based on *in vitro* concordance) was associated with a doubled adjusted odds of mortality (OR: 2.02, 95% CI: 1.86–2.20) [41]. In that review, the authors highlighted the large heterogeneity found across the studies and showed the variability in the pooled estimates for mortality in several subgroups. In the current systematic review, we have shown the diversity of nomenclature, definitions, and analytical methods in the underlying literature, and the high proportion of studies with medium or high risk of bias. Based on our findings, aggregation of data in meta-analysis will not provide better insights into the true association between inappropriate treatment and clinical outcomes among hospitalized patients with severe infections; it will rather provide a false sense of the strength of the evidence. To improve our understanding of the causative factors leading to increased mortality in patients with drug-resistant severe infections, exposure and outcome definitions need to be harmonized, and study quality needs to be improved. Recent efforts to promote uniformity across studies evaluating mortality in BSI patients identified and prioritized a minimum set of universal risk factors consistently associated with mortality in those patients [42].

Immortal time bias was an important concern in the reviewed studies, as time from admission-to-infection onset and time from infection to appropriate treatment were rarely addressed, potentially biasing estimates of the association between appropriate treatment and mortality [19]. In a similar manner, attributing total hospital LOS to the infection episode, thus also including the hospitalization days prior to the infection onset, has overestimated the association between inappropriate IAT and LOS [7,43]. Confounding by indication is another important problem; patients who present with severe illness are more likely to receive broad spectrum antibiotics, and hence, appropriate treatment. Although a substantial number of studies adjusted for the severity of acute illness, they frequently erroneously used postinfection parameters, without addressing the time-varying nature of this confounder. Adjustment for time-varying severity of illness is necessary to estimate causal effects, as the treating physicians might escalate antibiotic therapy in deteriorating patients, which some studies considered as part of IAT, but severity of illness is an intermediate variable on the causal pathway (i.e. severity of illness is affected by the history of antibiotic treatment). Using so-called g-methods (Robins' generalized methods), which are a set of advanced causal inference techniques (like G-computation, or marginal structural models), researchers can appropriately adjust for time-varying confounding while avoiding collider stratification bias [9,12]. However, none of the included studies applied these methods.

Studies assessing the association between inappropriate IAT and clinical outcomes among infected patients aim to understand the causal relationship. For a causal question, a causal framework is needed, which requires consistency (a well-defined (hypothetical) intervention), positivity (probability of receiving (mis) aligned IAT nonzero for every included patient), conditional exchangeability (no unmeasured confounding), and no interference (exposure of one individual should not affect potential outcome of another) [44]. In this systematic review, we have shown that there is a wide variability in IAT definitions, suggesting that the consistency assumption is frequently not fulfilled. Moreover, in the observational studies, any antibiotic treatment was considered for patients with and without inappropriate IAT, although the included antibiotics may not have been comparable in terms of efficacy and safety, and not all of these antibiotics would have been selected under trial conditions. In addition, patients with pan-drug-resistant BSIs, would have had zero probability of receiving appropriate IAT, discarding the positivity assumption. As discussed before, for most studies, we cannot assume conditional exchangeability. This indicates the clear need for guidelines to improve future studies.

Although nonrandomized studies cannot eliminate all confounding, applying robust methods in these studies can significantly reduce bias. One possible approach in this situation would be the use of a target trial framework, which compares exposures aligned with potential interventions and emulates such a trial using observational data [45]. This framework requires careful study planning, strict definitions for exposures and outcomes, and well-described analytical methods for the estimand of choice. In the IAT context, a potential intervention could be a rapid diagnostic test that reduces the delay to appropriate IAT, with the estimand of choice being excess 30-day mortality with post-discharge follow-up. Target trial emulation is increasingly applied in research on treatment duration and time to appropriate treatment initiation and could improve evidence quality and thus increase public health impact [46,47].

Our review is subject to some limitations, some of which are a direct result of the reviewed literature. First, most studies included a mix of hospital-acquired and community-acquired infections and did not focus specifically on one category. Each of these BSI categories ideally needs a different analytical approach, as time zero should be defined differently. Second, we included studies where treatment (in)appropriateness was not the primary study focus, as such methods may not have been optimal to assess its impact. Other limitations resulted from the review process; we focused on BSIs due to a specific subset of pathogens, although there is no reason to assume that findings would have been very different for another selection. Last, we did not perform double extraction, but we piloted the data extraction process as a team and held regular meetings to harmonize extraction strategies.

In conclusion, our findings highlight the need to standardize and refine criteria and methodologies for assessing the association between antibiotic treatment appropriateness and clinical outcomes to generate better evidence. Harmonization will also improve the comparability of studies and facilitate data aggregation in future systematic reviews. With more robust studies, the causal effect of (in)appropriate antibiotic treatment on patients' outcomes can be better understood and inform prioritization of public health policies.

CRediT authorship contribution statement

Marlieke E.A. de Kraker: Conceptualization: **Marlieke E.A. de Kraker, Nasreen Hassoun-Kheir, Alexander M. Aiken, Andrew J. Stewardson, Appiah-Korang Labi, Michael Loftus:** Methodology. **Nasreen Hassoun-Kheir:** Formal analysis. **Nasreen Hassoun-Kheir**, **Marlieke E.A. de Kraker**, **Alexander M. Aiken, Andrew J. Stewardson, Miriam Roncal Redin, Michael Loftus, Appiah-Korang Labi:** Investigation. **Nasreen Hassoun-Kheir and Miriam Roncal Redin:** Data curation. **Nasreen Hassoun-Kheir, Koen B. Pouwels, Marlieke E.A. de Kraker:** Writing – Original draft. Marlieke E.A. de Kraker: Supervision. All authors: Writing – Review & Editing.

Transparency declaration

Potential conflict of interest

N.H.K., S.H., and M.d.K. have received funding from the Innovative Medicines Initiative 2 Joint Undertaking under grant agreement number 101034420 (Predicting the Impact of Monoclonal Antibodies and Vaccines on Antimicrobial Resistance (PrIMAVeRa)). This Joint Undertaking receives support from the European Union's Horizon 2020 research and innovation program and European Federation of Pharmaceutical Industries and Associations (EFPIA). This communication reflects the author's view and neither Innovative Medicines Initiative nor the European Union, EFPIA, or any associated partners are responsible for any use that may be made of the information contained herein.

Financial report

This project has received funding from: the European Society of Clinical Microbiology and Infectious Diseases Young Investigator research grant (#18142) granted to N.H.K. in 2024; the Bill and Melinda Gates Foundation, grant reference number INV-003519 (Mortality from Bacterial Infections Resistant to Antibiotics (MBIRA)) for A.A., K.B.P., and M.S. are supported by the Antimicrobial Resistance, Prescribing, and Consumption Data to Inform Country Antibiotic Guidance and Local Action (ADILA) Project sponsored by the Wellcome Trust (222051/Z/20/Z).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cmi.2025.07.033.

References

[1] Hassoun-Kheir N, Guedes M, Ngo Nsoga MT, Argante L, Arieti F, Gladstone BP, et al. A systematic review on the excess health risk of antibiotic-resistant bloodstream infections for six key pathogens in Europe. Clin Microbiol Infect 2024;30:S14–25. https://doi.org/10.1016/j.cmi.2023.09.001.
[2] Kingston R, Vella V, Pouwels KB, Schmidt JE, El-Abasiri RAA, Reyna-Villasmil E, et al. Excess resource use and cost of drug-resistant infections for six key pathogens in Europe: a systematic review and Bayesian meta-analysis. Clin Microbiol Infect 2024;30:S26–36. https://doi.org/10.1016/j.cmi.2023.12.013.
[3] Bassetti M, Rello J, Blasi F, Goossens H, Sotgiu G, Tavoschi L, et al. Systematic review of the impact of appropriate versus inappropriate initial antibiotic therapy on outcomes of patients with severe bacterial infections. Int J Antimicrob Agents 2020;56:106184. https://doi.org/10.1016/j.ijantimicag.2020.106184.
[4] Paul M, Shani V, Muchtar E, Kariv G, Robenshtok E, Leibovici L. Systematic review and meta-analysis of the efficacy of appropriate empiric antibiotic therapy for sepsis. Antimicrob Agents Chemother 2010;54:4851–63. https://doi.org/10.1128/aac.00627-10.
[5] de Kraker MEA, Lipsitch M. Burden of antimicrobial resistance: Compared to what? Epidemiol Rev 2022;43:53–64. https://doi.org/10.1093/epirev/mxab001.
[6] Harbarth S, Nobre V, Pittet D. Does antibiotic selection impact patient outcome? Clin Infect Dis 2007;44:87–93. https://doi.org/10.1086/510075.
[7] Zhang Y, Du M, Johnston JM, Andres EB, Suo J, Yao H, et al. Estimating length of stay and inpatient charges attributable to hospital-acquired bloodstream infections. Antimicrob Resist Infect Control 2020;9:137. https://doi.org/10.1186/s13756-020-00796-5.
[8] Schuttevaer R, Alsma J, Brink A, van Dijk W, de Steenwinkel JEM, Lingsma HF, et al. Appropriate empirical antibiotic therapy and mortality: Conflicting data explained by residual confounding. PLOS One 2019;14:e0225478. https://doi.org/10.1371/journal.pone.0225478.
[9] Igelström E, Craig P, Lewsey J, Lynch J, Pearce A, Katikireddi SV. Causal inference and effect estimation using observational data. J Epidemiol Commun Health 2022;76:960. https://doi.org/10.1017/S0033291720005127.
[10] Cole SR, Frangakis CE. The consistency statement in causal inference: a definition or an assumption? Epidemiology 2009;20:3–5. https://doi.org/10.1097/EDE.0b013e31818ef366.
[11] VanderWeele TJ. On Well-defined hypothetical interventions in the potential outcomes framework. Epidemiology 2018;29:e24–5. https://doi.org/10.1097/EDE.0000000000000823.
[12] Pouwels KB, Vansteelandt S, Batra R, Edgeworth J, Wordsworth S, Robotham JV. Estimating the effect of healthcare-associated infections on excess length of hospital stay using inverse probability-weighted survival curves. Clin Infect Dis 2020;71:e415–20. https://doi.org/10.1093/cid/ciaa136.
[13] Karamitros G, Grant MP, Lamaris GA. Associations in medical research can be misleading: a clinician's guide to causal inference. J Surg Res 2025;310:145–54. https://doi.org/10.1016/j.jss.2025.03.043.
[14] Salas M, Hofman A, Stricker BH. Confounding by indication: an example of variation in the use of epidemiologic terminology. Am J Epidemiol 1999;149:981–3. https://doi.org/10.1093/oxfordjournals.aje.a009758.
[15] Ananth CV, Schisterman EF. Confounding, causality, and confusion: The role of intermediate variables in interpreting observational studies in obstetrics. Am J Obstet Gynecol 2017;217:167–75. https://doi.org/10.1016/j.ajog.2017.04.016.
[16] Lu H, Gonsalves GS, Westreich D. Selection bias requires selection: The case of collider stratification bias. Am J Epidemiol 2024;193:407–9. https://doi.org/10.1093/aje/kwad213.
[17] Tönnies T, Kahl S, Kuss O. Collider bias in observational studies. Dtsch Arztebl Int 2022;119:107–22. https://doi.org/10.3238/arztebl.m2022.0076.
[18] Lévesque LE, Hanley JA, Kezouh A, Suissa S. Problem of immortal time bias in cohort studies: Example using statins for preventing progression of diabetes. BMJ 2010;340:b5087. https://doi.org/10.1136/bmj.b5087.
[19] Yadav K, Lewis RJ. Immortal time bias in observational studies. JAMA 2021;325:686–7. https://doi.org/10.1001/jama.2020.9151.
[20] Hassoun-Kheir N, Labi KAK, Loftus M, Stewardson A, Aiken A, Pouwels K, et al. Nomenclature, definitions and methodological approaches to determine the impact of antimicrobial therapy on drug-resistant bloodstream infections PROSPERO International prospective register of systematic reviews. 2022. https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42022286778.
[21] Methley AM, Campbell S, Chew-Graham C, McNally R, Cheraghi-Sohi S. PICO, PICOS and SPIDER: A comparison study of specificity and sensitivity in three search tools for qualitative systematic reviews. BMC Health Serv Res 2014;14:579. https://doi.org/10.1186/s12913-014-0579-0.
[22] Pezzani MD, Tornimbene B, Pessoa-Silva C, de Kraker M, Rizzardo S, Salerno ND, et al. Methodological quality of studies evaluating the burden of drug-resistant infections in humans due to the WHO Global Antimicrobial Resistance Surveillance System target bacteria. Clin Microbiol Infect 2021;27:687–96. https://doi.org/10.1016/j.cmi.2021.01.004.
[23] Covidence. Covidence systematic review software. Melbourne: Veritas Health Innovation Melbourne. Available from: https://www.covidence.org

[24] JBI JBI. Critical appraisal tools. Adelaide: The University of Adelaide; 2018. Available from: https://jbi.global/critical-appraisal-tools.
[25] Lee Y, Lee YT, Wang YC, Chen CT, Sun JR, Liu CP, et al. Risk of mortality of catheter-related bloodstream infections caused by acinetobacter species: Is early removal of the catheters associated with a better survival outcome? J Inten Care Med 2018;33:361–9. https://doi.org/10.1177/0885066616677710.
[26] Thom KA, Shardell MD, Osih RB, Schweizer ML, Furuno JP, Perencevich EN, et al. Controlling for severity of illness in outcome studies involving infectious diseases: Impact of measurement at different time points. Infect Control Hosp Epidemiol 2008;29:1048–53. https://doi.org/10.1086/591453.
[27] Osih RB, McGregor JC, Rich SE, Moore AC, Furuno JP, Perencevich EN, et al. Impact of empiric antibiotic therapy on outcomes in patients with *Pseudomonas aeruginosa* bacteremia. Antimicrob Agents Chemother 2007;51:839–44. https://doi.org/10.1128/aac.00901-06.
[28] Rottier WC, Pinholt M, van der Bij AK, Arpi M, Blank SN, Nabuurs-Franssen MH, et al. Attributable mortality of vancomycin resistance in ampicillin-resistant Enterococcus faecium bacteremia in Denmark and the Netherlands: A matched cohort study. Infect Control Hosp Epidemiol 2022;43: 719–27. https://doi.org/10.1017/ice.2021.216.
[29] Nutman A, Glick R, Temkin E, Hoshen M, Edgar R, Braun T, et al. A case-control study to identify predictors of 14-day mortality following carbapenem-resistant Acinetobacter baumannii bacteraemia. Clin Microbiol Infect 2014;20:O1028–34. https://doi.org/10.1111/1469-0691.12716.
[30] Schweizer ML, Furuno JP, Harris AD, Johnson JK, Shardell MD, McGregor JC, et al. Empiric antibiotic therapy for Staphylococcus aureus bacteremia may not reduce in-hospital mortality: A retrospective cohort study. PLoS One 2010;5:e11432. https://doi.org/10.1371/journal.pone.0011432.
[31] Shi Q, Huang C, Xiao T, Wu Z, Xiao Y. A retrospective analysis of Pseudomonas aeruginosa bloodstream infections: Prevalence, risk factors, and outcome in carbapenem-susceptible and -non-susceptible infections. Antimicrob Resist Infect Control 2019;8:68. https://doi.org/10.1186/s13756-019-0520-8.
[32] Fitzpatrick JM, Biswas JS, Edgeworth JD, Islam J, Jenkins N, Judge R, et al. Gram-negative bacteraemia: A multi-centre prospective evaluation of empiric antibiotic therapy and outcome in English acute hospitals. Clin Microbiol Infect 2016;22:244–51. https://doi.org/10.1016/j.cmi.2015.10.034.
[33] Yilmaz M, Elaldi N, Balkan İ, Arslan F, Batırel AA, Bakıcı MZ, et al. Mortality predictors of Staphylococcus aureus bacteremia: A prospective multicenter study. Ann Clin Microbiol Antimicrob 2016;15:7. https://doi.org/10.1186/s12941-016-0122-8.
[34] Madney Y, Aboubakr S, Khedr R, Hafez H, Ahmed N, Elsheshtawy K, et al. Carbapenem-resistant enterobacteriaceae (CRE) among children with cancer: Predictors of mortality and treatment outcome. Antibiotics 2023;12:405. https://doi.org/10.3390/antibiotics12020405.
[35] Corcione S, Longo BM, Scabini S, Pivetta E, Curtoni A, Shbaklo N, et al. Risk factors for mortality in Acinetobacter baumannii bloodstream infections and development of a predictive mortality model. J Glob Antimicrob Resist 2024;38:317–26. https://doi.org/10.1016/j.jgar.2024.06.010.
[36] de Souza ILA, Cappellano P, Ferreira DB, Bergamasco MD, das Chagas Neto TC, Kerbauy FR, et al. Carbapenem-resistant Klebsiella pneumoniae bloodstream infections in haematological malignances and hematopoietic stem cell transplantation: Clinical impact of combination therapy in a 10-year Brazilian cohort. PLOS One 2024;19:e0297161. https://doi.org/10.1371/journal.pone.0297161.
[37] Nasomsong W, Changpradub D, Vasikasin V. Impact of inappropriate empirical antibiotic on outcomes in community-acquired third generation cephalosporin resistant enterobacterales bacteremia. Infect Chemother 2022;54:722–32. https://doi.org/10.3947/ic.2022.0096.
[38] Brigmon MM, Bookstaver PB, Kohn J, Albrecht H, Al-Hasan MN. Impact of fluoroquinolone resistance in Gram-negative bloodstream infections on healthcare utilization. Clin Microbiol Infect 2015;21:843–9. https://doi.org/10.1016/j.cmi.2015.05.013.
[39] Thom KA, Schweizer ML, Osih RB, McGregor JC, Furuno JP, Perencevich EN, et al. Impact of empiric antimicrobial therapy on outcomes in patients with *Escherichia coli* and *Klebsiella pneumoniae* bacteremia: A cohort study. BMC Infect Dis 2008;8:116. https://doi.org/10.1186/1471-2334-8-116.
[40] Tumbarello M, Sali M, Trecarichi EM, Leone F, Rossi M, Fiori B, et al. Bloodstream infections caused by extended-spectrum-beta-lactamase- producing *Escherichia coli*: Risk factors for inadequate initial antimicrobial therapy. Antimicrob Agents Chemother 2008;52:3244–52. https://doi.org/10.1128/AAC.00063-08.
[41] Hung YP, Lee CC, Ko WC. Effects of inappropriate administration of empirical antibiotics on mortality in adults with bacteraemia: Systematic review and meta-analysis. Front Med (Lausanne) 2022;9:869822. https://doi.org/10.3389/fmed.2022.869822.
[42] Varon B, Palacios-Baena ZR, de Kraker MEA, Rodríguez-Baño J, Leibovici L, Paul M, et al. Universal Risk Factors for Mortality in Bloodstream Infections (UNIFORM): A systematic review and Delphi survey. Clin Microbiol Infect 2024;30:453–61. https://doi.org/10.1016/j.cmi.2023.12.030.
[43] Stewart S, Robertson C, Pan J, Kennedy S, Haahr L, Manoukian S, et al. Impact of healthcare-associated infection on length of stay. J Hosp Infect 2021;114: 23–31. https://doi.org/10.1016/j.jhin.2021.02.026.
[44] Rehkopf DH, Glymour MM, Osypuk TL. The consistency assumption for causal inference in social epidemiology: When a rose is not a rose. Curr Epidemiol Rep 2016;3:63–71. https://doi.org/10.1007/s40471-016-0069-5.
[45] Hernán MA, Wang W, Leaf DE. Target trial emulation: A framework for causal inference from observational data. JAMA 2022;328:2446–7. https://doi.org/10.1001/jama.2022.21383.
[46] Smith LH, García-Albéniz X, Chan JM, Zhao S, Cowan JE, Broering JM, et al. Emulation of a target trial with sustained treatment strategies: An application to prostate cancer using both inverse probability weighting and the g-formula. Eur J Epidemiol 2022;37:1205–13. https://doi.org/10.1007/s10654-022-00929-7.
[47] Lim C, Mo Y, Teparrukkul P, Hongsuwan M, Day NPJ, Limmathurotsakul D, et al. Effect of delays in concordant antibiotic treatment on mortality in patients with hospital-acquired acinetobacter species bacteremia: Emulating a target randomized trial with a 13-year retrospective cohort. Am J Epidemiol 2021;190:2395–404. https://doi.org/10.1093/aje/kwab158.