Clinical Epidemiology

Dovepress
open access to scientific and medical research

Open Access Full Text Article

ORIGINAL RESEARCH

Antimicrobial Resistance and Mortality in Hospitalized Patients with Bacteremia in the Greater Paris Area from 2016 to 2019

Salam Abbara[1,2], Didier Guillemot[1–3], Salma El Oualydy[4], Maeva Kos[4], Cécile Poret[5], Stéphane Breant[5], Christian Brun-Buisson[1,2], Laurence Watier[1,2]

[1]Anti-Infective Evasion and Pharmacoepidemiology Team, Inserm, UVSQ, University Paris-Saclay, CESP, Montigny-Le-Bretonneux, France; [2]Institut Pasteur, Epidemiology and Modelling of Antibiotic Evasion (EMAE), University Paris Cité, Paris, France; [3]Public Health, Medical Information, Clinical Research, AP-HP, University Paris Saclay, Le Kremlin-Bicêtre, France; [4]Plateforme des données de santé - Health Data Hub, Paris, France; [5]AP-HP, Direction des Systèmes d'Information, Pôle Innovation et Données, Paris, France

Correspondence: Salam Abbara, Institut Pasteur, Epidemiology and Modelling of Antibiotic Evasion (EMAE) Unit, 25-28 rue du Docteur Roux, Paris, 75015, France, Tel +33 1 45 68 80 00, Email salam.abbara@gmail.com

Purpose: Antibiotic-resistant bacteremia is a leading global cause of infectious disease morbidity and mortality. Clinical data warehouses (CDWs) allow for the secure, real-time coupling of diverse data sources from real-world clinical settings, including care-based medical-administrative data and laboratory-based microbiological data. The main purpose of this study was to assess the contribution of CDWs in the epidemiological study of antibiotic resistance by constructing a database of bacteremia patients, BactHub, and describing their main clinico-microbiological features and outcomes.

Patients and Methods: Adult patients with bacteremia hospitalized between January 1, 2016 and December 31, 2019 in 14 acute care university hospitals from the Greater Paris area were identified; their first bacteremia episode was included. Data describing patients, episodes of bacteremia, bacterial isolates, and antimicrobial resistance were structured.

Results: Among 29,228 patients with bacteremia, 41% of episodes were community-onset (CO) and 59% were hospital-acquired (HA). Thirty-day and ninety-day mortality rates were 15% and 20% in CO episodes, and 18% and 36% in HA episodes. Overall resistance rates were high, including third-generation cephalosporin resistance among *Klebsiella pneumoniae* (CO 21%, HA 37%) and *Escherichia coli* (CO 13%, HA 17%), and methicillin resistance among *Staphylococcus aureus* (CO 11%, HA 14%). Annual incidence rates increased significantly from 2017 to 2019, from 20.0 to 20.9 to 22.1 stays with bacteremia per 1000 stays ($p < 0.0001$).

Conclusion: The Bacthub database provides accurate clinico-microbiological data describing bacteremia across France's largest hospital group. Data from Bacthub may inform surveillance and the clinical decision-making process for bacteremia patients, including choice of antimicrobial therapy. The database also offers opportunities for research, including analysis of hospital care pathways and significant patient outcomes such as mortality and recurrence of infection.

Keywords: data warehousing, bacteremia, drug resistance, microbial, anti-bacterial agents, mortality, incidence

Introduction

Infections due to antibiotic-resistant bacteria are a leading cause of morbidity and mortality.[1,2] A thorough understanding of their epidemiology requires comprehensive study of their various characteristics. These include clinical characteristics such as the patient's diagnosis, treatment, course, infection severity, outcomes and follow-up; as well as microbiological characteristics such as the bacterial isolate's species and antibiotic sensitivity. Yet integrating such diverse characteristics into a single database for epidemiological study is a great challenge. While many large-scale, highly detailed microbiological databases exist, few include clinical data, mainly due to difficulty in combining both types of data on a large scale. On the other hand, administrative databases usually lack key clinical and microbiological data needed to assess bacterial resistance and measure infection severity.[3]

Received: 15 August 2022
Accepted: 12 November 2022
Published: 14 December 2022

Clinical data warehouses (CDWs) help to address this challenge by providing access to multiple large-scale data sources from real-world clinical settings.[4–6] When properly qualified and structured, they can include large numbers of patients from diverse healthcare institutions without a need to access their individual medical files, and can be used for surveillance or to conduct investigative, evaluative or observational studies at a (multi-)hospital level.[5–9] In particular, by allowing simultaneous access to both clinical and microbiological data, CDWs have the potential to help provide more comprehensive epidemiological descriptions of acute bacterial infections across large patient populations over time.[4,10] Bacteremia is a severe acute bacterial infection which is associated with significant morbidity and mortality, and high levels of antimicrobial resistance (AMR).[1] While the diagnosis of bacteremia is straightforward based solely on microbiological data, requiring only the exclusion of potential contaminants, clinical data are crucial to fully understand bacteremia epidemiology, including impacts on human health.

In this study, we aimed to assess the contribution of CDWs in the study of bacteremia by constructing a database of patients hospitalized with bacteremia and describing the main clinico-microbiological features and outcomes of included infections. The Bacthub database was thus established, based on the CDW of the university hospital group of the Greater Paris area, Assistance Publique - Hôpitaux de Paris (AP-HP).[10] Here, we report on the selection of the patient population; describe the main characteristics of the included patients, bacteremia episodes and bacterial isolates; and provide an incidence estimate over three years. Finally, we discuss the representativeness and usefulness of the database and next steps for expanding the use of CDWs to facilitate the epidemiological study of AMR.

Materials and Methods

Source of the Data

Data were obtained retrospectively from the AP-HP CDW (Supplementary File S1.1).[10] The AP-HP (https://www.aphp.fr) is the largest university hospital entity in Europe, with 39 hospitals (22,474 beds) located mainly in the region of Île-de-France (the Greater Paris area) and totaling 1.5 million hospitalizations per year (10% of all hospitalizations in France).

Study Population

Hospital stays of adults (≥ 18 years) with at least one positive blood culture between January 1, 2016 and December 31, 2019 in any of the 23 AP-HP hospitals offering acute care were selected. Pediatric hospitals were excluded. Only stays > 24h were retained, unless the patient died within 24h. Only the first hospital stay of each patient over the 4-year study period was retained, and only the first bacteremia episode of the stay was analyzed. Episodes could include multiple positive blood cultures provided that each was collected less than 72 hours after the previous one. Contaminants were excluded. Species classified as contaminants were *Corynebacterium sp., Staphylococcus epidermidis, Staphylococcus pettenkoferi, Staphylococcus haemolyticus, Staphylococcus hominis, coagulase-negative Staphylococci, Staphylococcus capitis, Staphylococcus hominis subsp. hominis, Micrococcus luteus, Staphylococcus capitis subsp. capitis*, or *Staphylococcus warneri*. Within the same episode, we considered a blood culture positive for one of these species not to be a contaminant if: 1) there was at least one other blood culture positive with the same species within 72h, AND 2) both blood cultures had at least 6 out of 8 Antibiotic susceptibility testing (AST) results available for Oxacillin, Cefoxitin, Tobramycin, Gentamicin, Rifampicin, Erythromycin, Clindamycin, Cotrimoxazole, AND 3) available AST results for these antibiotics were similar between both blood cultures. Finally, to avoid selection and temporal bias, we excluded hospitals with incomplete or unstable microbiology reports over the study period (Supplementary File S1.2).

Data Collection

For each patient, gender, age, diagnoses (International Classification of Diseases (ICD-10) codes) and death date (if any) were extracted from the AP-HP CDW. For each episode, admission and discharge dates, intensive care unit (ICU) stay, date of blood culture collection, identified bacteria were extracted. For each bacterial isolate, antibiotic susceptibility testing (AST) results were extracted. Reported AST results used EUCAST clinical breakpoints (https://eucast.org/clinical_breakpoints/) and the qualitative susceptibility categories of "Susceptible, standard dosing regimen" (S), "Susceptible, increased dosage" (I) and "Resistant" (R).

Variables

Variables included in our analysis were either extracted directly from the database (eg gender) or calculated using extracted data (eg Charlson comorbidity index calculated using comorbidity-associated ICD-10 codes).

Patient variables included were gender, age, Charlson comorbidity index,[11] ICD-10 codes for comorbidities,[11] and mortality (in-hospital, in-hospital with or without ICU admission, 30-day and 90-day after first positive blood culture).

Episode variables included were the number of polymicrobial episodes, the length of stay (LOS) with bacteremia (defined as the time from the first positive blood culture to hospital discharge), ICU admission, and the primary site of infection (absence, cardio-vascular, congenital-perinatal, device-related, digestive tract, eye ear nose and throat, genital, bone and joint, lower respiratory tract, muscle, nervous system, sexually-transmitted, skin and soft tissues, urinary tract, others, and multiple primary sites; defined according to ICD-10 codes, as detailed in Supplementary Table S1).[3] Bacteremia episodes were classified as primary or secondary based on the absence or presence, respectively, of a recorded primary site. Episodes were classified as community-onset (CO) if the first positive blood culture was sampled < 48 hours after hospital admission, excluding patients transferred or discharged from an AP-HP hospital within 7 days. Other episodes were classified as hospital-acquired (HA).

Bacterial isolate variables included were qualitative AST results for antibiotics commonly used to treat infections caused by that bacterial species. These were analyzed after deduplication by species and episode, favoring the most resistant AST results. For all antibiotics tested, except gentamicin in enterococci, strains reported as "susceptible, increased exposure" (I) were considered resistant. AMR rates were assessed according to the European Centre for Disease Prevention and Control (ECDC) combinations.[1]

Statistical Analysis

Patients, episodes, and bacterial isolates were analyzed separately, stratified by category of bacteremia (CO vs HA). Patient, episode, and bacterial variables are expressed as the median [first quartile (Q1), third quartile (Q3)] for continuous variables, and as counts (%) for categorical variables. Data are reported for the whole 4-year study period, and patient variables are stratified by year in the Supplementary Material. AMR rates were compared to ECDC data (based on blood or cerebrospinal fluid isolates) for France, which are reported to have a high representativeness.[1]

Annual bacteremia incidence was calculated from 2017 to 2019 and expressed as the total number of stays with bacteremia per 1000 stays. The total number of stays was collected from French Annual Health Facilities Statistics data.[12] Incident stays were selected by excluding stays with a previous hospitalization with bacteremia within the past 12 months. Changes in incidence were assessed by a $\chi 2$ test for trend, and a p value <0.05 was considered significant.

HiveQL, Python 3 and Spark Python 2.4.3 programming languages were used to structure the database and to compute statistical analyses. This study follows the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) reporting guideline.

Results

From 2016 to 2019, we identified 30,877 patients with clinically significant episodes of bacteremia in 23 hospitals. After excluding pediatric hospitals and those failing to meet quality checks, data from 14 hospitals were retained (Flowchart, Figure 1; Supplementary Table S2), covering approximately 22% of short stays in the Île-de-France region. Finally, a total of 29,228 patients (95%), with 29,228 first episodes of bacteremia (CO 41%, HA 59%) were included.

Patient Characteristics

Patients with a CO episode had a slightly higher median age at infection (68 vs 64 years), with a higher proportion of elderly patients (>80 years; 23% vs 13%), and fewer men (55% vs 62%) (Table 1). They had fewer comorbidities (Charlson score >0, 61% vs 78%), including half as many cancers (25% vs 43%), metastatic solid tumors (10% vs 16%) and vascular diseases (12% vs 18%). In comparison, patients with a HA episode had higher mortality rates (in-hospital, 19% vs 14%; 30-day, 18% vs 15%; 90-day, 26% vs 20%). Notably, the 30-day mortality rate was slightly lower than the in-hospital mortality rate in the HA group (18% vs 19%), possibly due to stays longer than 30 days. In both groups, ICU

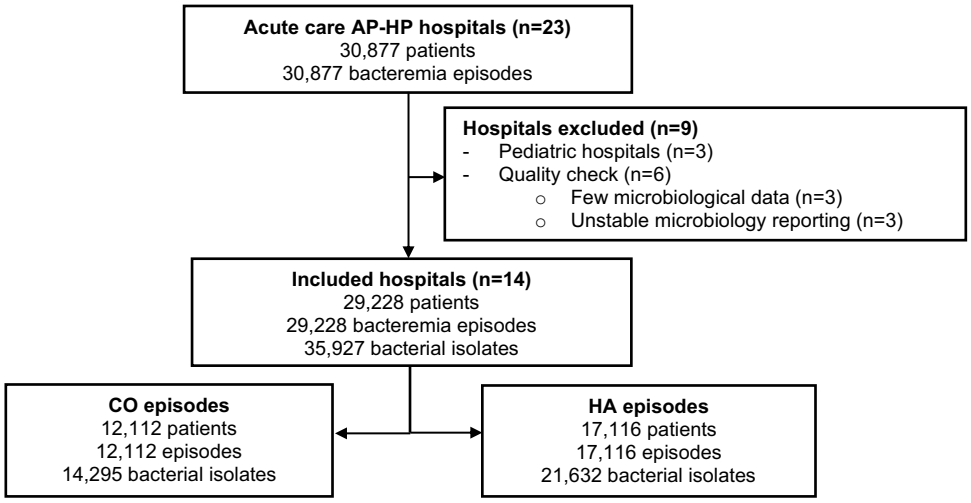

Figure 1 Flowchart of the selection process for the Bacthub database, Greater Paris university hospitals, 2016–2019.
Abbreviations: AP-HP, Assistance Publique - Hôpitaux de Paris; CO, community onset; HA, hospital acquired.

admission was associated with high in-hospital, 30-day and 90-day mortality rates (26–28%, 25–24%, 31–32%, respectively) compared to patients without ICU admission (8–13%, 10–13%, 15–22%, respectively). Patients' characteristics were stable between 2016 and 2019 (Supplementary Table S3).

Episode Characteristics

The median LOS with bacteremia was markedly higher in HA episodes (15 vs 8 days). ICU admission (45% vs 31%), polymicrobial bacteremia (19% vs 14%) and primary bacteremia (25% vs 22%) were also more frequent among HA episodes

Table 1 Characteristics of Patients with Bacteremia in the Bacthub Database, 2016–2019

	CO	HA
No. of patients	12,112 (41.44)	17,116 (58.66)
Men	6678 (55.14)	10,576 (61.79)
Age, years	68 [54–80]	64 [52–74]
18–35	953 (7.87)	1473 (8.61)
35–50	1496 (12.35)	2453 (14.33)
50–65	3002 (24.79)	5312 (31.04)
65–80	3823 (31.56)	5607 (32.76)
>80	2838 (23.43)	2271 (13.27)
Charlson score[a]	2 [0–3]	2 [1–5]
0	4159 (39.13)	3261 (22.18)
1–2	3418 (32.16)	5208 (35.42)
3–4	1331 (12.52)	2404 (16.35)
5–6	520 (4.89)	1059 (7.20)
7–8	870 (8.19)	1948 (13.25)
>8	330 (3.11)	822 (5.59)

(*Continued*)

Table 1 (Continued).

	CO	HA
Comorbidity[a]		
Any malignancy[b]	2709 (25.49)	6264 (42.61)
Metastatic solid tumor	1054 (9.92)	2363 (16.07)
Congestive heart failure	1572 (14.79)	2815 (19.15)
Diabetes	2102 (19.78)	2940 (20.00)
Uncomplicated	1315 (12.37)	1764 (12.00)
Chronic complication	787 (7.40)	1176 (8.00)
Vascular disease	1261 (11.86)	2716 (18.47)
Coronary	185 (1.74)	496 (3.37)
Peripheral	597 (5.62)	1237 (8.41)
Cerebrovascular	581 (5.47)	1347 (9.16)
Renal disease	1337 (12.58)	2152 (14.64)
Liver disease	1004 (9.45)	1757 (11.95)
Mild	659 (6.20)	945 (6.43)
Moderate / severe	345 (3.25)	812 (5.52)
Chronic pulmonary disease	716 (6.74)	1177 (8.01)
Dementia	532 (5.01)	562 (3.82)
Hemiplegia or paraplegia	296 (2.79)	941 (6.40)
Rheumatologic disease	154 (1.45)	240 (1.63)
Mortality		
In-hospital	1636 (13.51)	3302 (19.29)
ICU admission	993/3758 (26.42)	2110/7615 (27.71)
No ICU admission	643/8354 (7.70)	1192/9501 (12.55)
30-day	1765 (14.57)	3028 (17.69)
ICU admission	951/3758 (25.31)	1788/7615 (23.48)
No ICU admission	814/8354 (9.74)	1240/9501 (13.05)
90-day	2376 (19.62)	4451 (26.00)
ICU admission	1154/3758 (30.71)	2411/7615 (31.66)
No ICU admission	1222/8354 (14.63)	2040/9501 (21.47)

Notes: Data are expressed as the median [first quartile, third quartile] for continuous variables, and as counts (%) for categorical variables. [a]Data available for 10,628 and 14,702 patients in the CO and HA group, respectively. [b]Including leukemia and lymphoma.
Abbreviations: CO, community-onset; HA, hospital-acquired; ICU, intensive care unit.

(Table 2). In secondary bacteremia, multiple primary sites of infection were frequently identified, especially in HA episodes (35% vs 24%). The main unique primary sites of infection were the urinary tract (UTI, CO 19%, HA 9%), the lower respiratory tract (LRTI, CO 11%, HA 10%), and the digestive tract (DTI, CO 9%, HA 6%), with a clear predominance of UTI in CO episodes. Device-related infections accounted for 6–7% of secondary infections in both groups, among which venous or arterial catheter-related infections were predominant (3% and 5% in the CO and HA groups, respectively). Skin and soft tissue, cardio-vascular and bone-joint infections were less frequent in both groups (3–4%, 1–2% and 1%, respectively).

Bacterial Characteristics and Antimicrobial Resistance

After deduplication by species and episode, 35,927 bacterial isolates were analyzed. The three most identified species (Table 3) among both CO and HA episodes were *Escherichia coli* (CO 27%, HA 15%), *Staphylococcus aureus* (12% in

Table 2 Clinical Features of Bacteremia in the Bacthub Database, 2016–2019

	CO	HA
No. of episodes	12,112 (41.44)	17,116 (58.66)
Polymicrobial	1660 (13.71)	3318 (19.39)
LOS with bacteremia, days	8.0 [3.0–17.0]	15.0 [7.0–29.0]
≤7	5903 (48.7)	4592 (26.8)
7–14	2599 (21.46)	3908 (22.83)
14–30	2350 (19.40)	4663 (27.24)
30–60	992 (8.19)	2669 (15.59)
>60	268 (2.21)	1284 (7.50)
ICU admission	3758 (31.03)	7615 (44.49)
Infection site[a,b]		
Primary bacteremia[c]	2292 (21.57)	3606 (24.53)
Secondary bacteremia		
Multiple sites[d]	2544 (23.94)	5202 (35.38)
Urinary tract	2034 (19.14)	1252 (8.52)
Lower respiratory tract	1151 (10.83)	1511 (10.28)
Digestive tract	968 (9.11)	839 (5.71)
Device-related	594 (5.59)	1054 (7.17)
Skin and soft tissue	386 (3.63)	493 (3.35)
Cardio-vascular	178 (1.68)	107 (0.73)
Bone and Joint	155 (1.46)	119 (0.81)
Others[e]	109 (1.03)	316 (2.15)

Notes: Data are expressed as the median [first quartile, third quartile] for continuous variables, and as counts (%) for categorical variables. [a]Only the top 10 primary infection sites in each category are shown. [b]Data available for 10,628 and 14,702 episodes of CO and HA bacteremia, respectively. [c]No ICD-10 code of bacterial infection during the hospital stay, excluding bacteremia codes. [d]Presence of ICD-10 codes referring to multiple categories of bacterial infection. [e]This category is mainly composed of codes T81.4 (Infection following a procedure), A49 (Bacterial infection of unspecified site) and R02 (Gangrene, not elsewhere classified).

Abbreviations: CO, community-onset; HA, hospital-acquired; ICU, intensive care unit; LOS, length of stay.

Table 3 Distribution of Bacterial Isolates in the Bacthub Database, 2016–2019

	CO	HA
No. of unique bacterial species	14,295 (39.79)	21,632 (60.21)
E. coli	3832 (26.81)	3151 (14.57)
S. aureus	1672 (11.70)	2513 (11.62)
K. pneumoniae	754 (5.28)	1500 (6.93)
S. pneumoniae	529 (3.70)	–
S. epidermidis	354 (2.48)	1403 (6.49)
P. aeruginosa	345 (2.41)	1427 (6.60)
E. faecalis	448 (3.13)	1069 (4.94)
P. mirabilis	225 (1.57)	–
S. agalactiae	215 (1.50)	–
S. pyogenes	192 (1.34)	–
E. cloacae	–	832 (3.85)
E. faecium	–	627 (2.90)
B. fragilis	–	346 (1.60)
S. haemolyticus	–	320 (1.48)

Notes: Bacterial isolates data are expressed as numbers (%). Only the 10 most identified species of each category are reported.

Abbreviations: *B. fragilis, Bacteroides fragilis*; CO, community-onset; *E. cloacae, Enterobacter cloacae; E. faecalis, Enterococcus faecalis; E. faecium, Enterococcus faecium; E. coli, Escherichia coli*; HA, hospital-acquired; *K. oxytoca, Klebsiella oxytoca; K. pneumoniae, Klebsiella pneumoniae; P. mirabilis, Proteus mirabilis; P. aeruginosa, Pseudomonas aeruginosa; S. aureus, Staphylococcus aureus; S. epidermidis, Staphylococcus epidermidis; S. haemolyticus, Staphylococcus haemolyticus; S. agalactiae, Streptococcus agalactiae* (group B); *S. dysgalactiae, Streptococcus dysgalactiae; S. pneumoniae, Streptococcus pneumoniae; S. pyogenes, Streptococcus pyogenes* (group A).

both groups) and *Klebsiella pneumoniae* (CO 5%, HA 7%). The fourth and fifth most identified species were *Streptococcus pneumoniae* (4%) and *Enterococcus faecalis* (3%) in CO episodes, and *Pseudomonas aeruginosa* (7%) and *Staphylococcus epidermidis* (7%) in HA episodes.

Mortality rates differed by pathogen and by bacteremia onset (Supplementary Table S4). The lowest mortality rates were found in CO bacteremia due to a *Streptococcus* species (*S. pneumoniae, S. agalactiae, S. pyogenes*) (in-hospital, 9–13%; 90-day, 14–16%) or in *Bacteroides fragilis* HA bacteremia (in-hospital, 14%; 90-day, 18%). Other etiologies, including *E. faecalis, S. aureus, S. epidermidis, K. pneumoniae, Proteus mirabilis, Enterobacter cloacae*, were associated with high in-hospital mortality rates (14–21%), which increased by 6 to 10% at 90 days (21–28%). For all bacteria studied in both the CO and HA groups, except *S. aureus* and *P. aeruginosa*, in-hospital and 90-day mortality rates were higher in the HA group. Bacteria associated with the highest mortality rates were *P. aeruginosa* (in-hospital, 24–25%; 90-day, 31–33%) in the CO and HA group, and especially *E. faecium* in the HA group (in-hospital, 40%; 90-day, 46%).

Third-generation cephalosporin resistance (3GC-R) was very frequent in *E. coli* and especially in *K. pneumoniae* isolates, in both CO (*E. coli*, 13%; *K. pneumoniae*, 21%) and HA episodes (*E. coli*, 17%; *K. pneumoniae*, 37%, Table 4). 3GC-R was more frequent in *K. pneumoniae* than *E. coli* in CO episodes only when the primary site was urinary (31% vs 14% for *E. coli*, Figure 2), but in HA episodes whatever the primary site. The rate of methicillin resistance in *S. aureus* isolates was close to 10–15% in both CO and HA episodes and regardless of the primary infection site, except for UTI in which case methicillin resistance reached 20–25%.

Table 4 Rates of Resistance to Antibiotics (in Percent) of Bacterial Isolates in the Bacthub Database, 2016–2019

	CO	HA
Escherichia coli		
Aminopenicillins	62.0	68.6
3GC	13.1	17.2
Carbapenems	0.1	0.4
Fluoroquinolones	21.9	23.0
Aminoglycosides	10.6	12.9
Staphylococcus aureus		
Methicillin	11.2	14.2
Rifampicin	1.4	1.7
Fluoroquinolones	10.1	12.3
Klebsiella pneumoniae		
3GC	21.0	37.1
Carbapenems	0.4	1.9
Fluoroquinolones	21.6	33.2
Aminoglycosides	15.3	27.1
Streptococcus pneumoniae		
Penicillin non-wild-type[a]	18.7	-
Macrolides	19.0	-
Fluoroquinolones	9.1	-
3GC	0.9	-
Staphylococcus epidermidis		
Methicillin	-	84.8
Rifampicin	-	16.2
Fluoroquinolones	-	70.7
Pseudomonas aeruginosa		
Pip-tazobactam	-	16.7
Ceftazidime	-	11.2
Carbapenems	-	20.4
Fluoroquinolones	-	12.1
Aminoglycosides[b]	-	6.8
Multi-drug resistance[c]		9.4

(*Continued*)

Table 4 (Continued).

	CO	HA
Enterococcus faecalis		
Gentamicin high-level resistance	8.1	-
Vancomycin	0.2	-
Enterococcus faecium		
Vancomycin		1.3

Notes: Resistance rates are expressed as percent. They are shown for the 5 most identified species of each category, and for *Enterococcus faecium*. For all antibiotics tested, except gentamicin in enterococci, strains reported as "susceptible, increased exposure" (I) were considered resistant. [a]Priority was given to penicillin susceptibility test results over oxacillin results. The term penicillin non-wild-type refers to *S. pneumoniae* isolates reported by local laboratories as "susceptible, increased exposure" (I) or resistant (R) to penicillin. [b]Tobramycin resistance. [c]Combined resistance to ≥3 antimicrobial groups among piperacillin+tazobactam, ceftazidime, carbapenems, fluoroquinolones and aminoglycosides, according to the definition of the ECDC.
Abbreviations: 3GC, Third generation cephalosporins; CO, community-onset; HA, hospital-acquired; HL, high-level; pip-tazobactam, piperacillin + tazobactam.

In CO episodes, 19% of *S. pneumoniae* isolates were penicillin non-wild-type, while 8% of *E. faecalis* isolates had a high-level resistance to gentamicin. In HA episodes, *Pseudomonas aeruginosa* isolates had significant levels of resistance to carbapenems (20%), piperacillin-tazobactam (17%), fluoroquinolones (12%), ceftazidime (11%), and multi-drug resistance (9%); the rate of vancomycin-resistant *E. faecium* was 1.3%.

Combining CO and HA episodes, *E. coli* resistance rates were somewhat higher than those reported by ECDC for France for all evaluated resistances (3GC, 15% vs 9–11%; fluoroquinolones, 22% vs 16–17%; aminoglycosides, 12% vs 7–8%), and were closer to overall Europe-wide ECDC rates (Tables 5; Supplementary Table S5).[1] By contrast, the MRSA rate (13%), and *K. pneumoniae* resistance rates (3GC, 32%; fluoroquinolones, 29%; aminoglycosides, 23%), were close to French ECDC rates. *P. aeruginosa* resistance rates (piperacillin-tazobactam, 15%; fluoroquinolones, 12%; ceftazidime, 11%; multi-drug resistance 9%) were also close to French ECDC rates, except for carbapenems (20% vs 13–16%) and aminoglycosides (7% vs 8–11%). Rates of high-level resistance to gentamicin among *E. faecalis* isolates (9% vs 12–15%) and non-wild-type penicillin phenotype among *S. pneumoniae* isolates (21% vs 25%) were lower in our data than in French ECDC data.

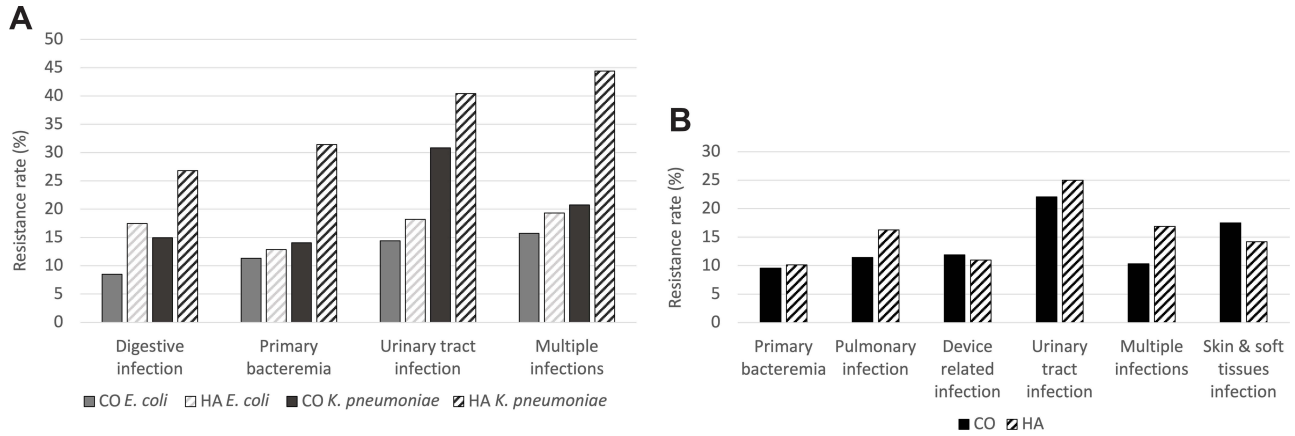

Figure 2 Rates of resistance of bacterial isolates in the blood to major antimicrobials, per primary infection site. (**A**) Third generation cephalosporins resistance rate of *E. coli* and *K. pneumoniae* isolates. (**B**) Methicillin resistance rate of *S. aureus* isolates.
Note: Only primary infection sites with at least 50 isolates per bacterial species are described.
Abbreviations: CO, community onset; HA, hospital acquired.

Table 5 Rates of Resistance to Antibiotics and Trend of Evolution in the Bacthub Database, Compared to ECDC Data

	All Years Bacthub (2016–2019)	Trend 2016–2019	All Years ECDC France (2016–2019)	Trend 2015–2019	All Countries ECDC (2016–2019)	Trend 2015–2019
E. coli						
Aminopenicillins	64.94%		57.2% - 54.5%	<	59.0% - 57.1%	<
3GC	14.95%	<	11.2% - 8.8%	<	14.9% - 15.1%	>
Carbapenems	0.25%	>	0% - 0%		0.1% - 0.3%	>
Fluoroquinolones	22.42%	<	16.7% - 16.0%	<	25.2% - 23.8%	<
Aminoglycosides	11.60%	<	7.9% - 7.0%	<	11.6% - 10.8%	<
S. aureus						
Methicillin	12.98%		13.8% - 11.6%	<	17.7% - 15.5%	<
K. pneumoniae						
3GC	31.70%		28.9% - 30.2%		31.4% - 31.3%	
Carbapenems	1.39%		0.4% - 1.0%	>	7.4% - 7.9%	>
Fluoroquinolones	29.30%		27.7% - 30.9%		30.3% - 31.2%	>
Aminoglycosides	23.19%		26.2% - 23.4%	<	24.4% - 22.3%	<
P. aeruginosa						
Pip-tazobactam	15.37%		16.0% - 16.7%	>	17.5% - 16.9%	<
Ceftazidime	10.47%		11.3% - 11.5%		14.4% - 14.3%	<
Carbapenems	19.86%		15.6% - 12.7%	<	18.2% - 16.5%	<
Fluoroquinolones	12.06%	<	13.6% - 13.7%	<	18.8% - 18.9%	<
Aminoglycosides[a]	6.48%	<	10.7% - 7.8%	<	14.0% - 11.5%	<
Multi-drug resistance[c]	9.00%	<	10.3% - 8.0%	<	13.4% - 12.1%	<
S. pneumoniae						
Penicillin NWT[b]	21.13%		25.3% - 25.3%		13.1% - 12.1%	<
Macrolides	22.00%		22.9% - 19.4%	<	16.6% - 14.5%	<
E. faecalis						
Gentamicin HLR	9.27%		15.0% - 12.0%	<	31.8% - 26.6%	<

Notes: For all antibiotics tested, except gentamicin in enterococci, strains reported as "susceptible, increased exposure" (I) were considered resistant. The statistical significance of trends was assessed by a $\chi 2$ test for trend, and a p-value of <0.05 was considered significant. Only statistically significant results are described (<, decreasing trend; >, increasing trend). [a]Tobramycin resistance rates in *Pseudomonas* isolates from the Bacthub database were compared to Tobramycin/Gentamicin/Netilmicin resistance rates in *Pseudomonas* isolates from the ECDC. [b]Priority was given to penicillin susceptibility test results over oxacillin results. The term penicillin non-wild-type refers to *S. pneumoniae* isolates reported by local laboratories as "susceptible, increased exposure" (I) or resistant (R) to penicillin. [c]Combined resistance to ≥3 antimicrobial groups among piperacillin+tazobactam, ceftazidime, carbapenems, fluoroquinolones and aminoglycosides, according to the definition of the ECDC.
Abbreviations: 3GC, Third generation cephalosporins; HLR, high-level resistance; NWT, non-wild-type; pip-tazobactam, piperacillin + tazobactam.

Incidence

Annual bacteremia incidence rates were 20.0 stays with bacteremia per 1000 stays in 2017 (7288/364,021 95% confidence interval (CI) 19.6–20.5), 20.9 in 2018 (7531/359,534, 95% CI 20.5–21.4), and 22.1 in 2019 (7767 / 351,367, 95% CI 21.6–22.6). Bacteremia incidence increased significantly from 2017 to 2019 ($p < 0.0001$).

Discussion

In this study, we describe the clinical and microbiological characteristics of almost 30,000 adult patients hospitalized with bacteremia between 2016 and 2019 across 14 acute care university hospitals in the Greater Paris area. The proportions of CO and HA episodes (41% and 59%) are consistent with estimates from the literature.[13–15] Among all episodes, 17% were polymicrobial; this rate appears to be approximately twice as high as in previous works.[13,14,16,17] In-hospital and 30-day mortality rates were similar, reaching 14–15% in patients with a CO episode, and 18–19% in patients with a HA episode. The 30-day mortality rate was similar in CO episodes, but lower in HA episodes, relative to estimates reported by two Danish studies from a decade ago.[18,19] Of note, our results show a marked difference in in-hospital and 30-day mortality rates between patients with (range, 24–28%) and without (8–13%) ICU admission. To our knowledge, no study has estimated the 90-day mortality of patients by bacteremia onset, regardless of the bacterial species. We found high 90-day mortality rates in both the CO (20%) and HA (26%) groups; these rates reached 31–32% in patients who were admitted to the ICU during their stay, compared to 15–22% among patients without an ICU stay. While all bacteria species were associated with high 90-day mortality rates, the highest rates were found for *P. aeruginosa* (~30%) and *E. faecium* (46%). Such high rates may be explained in part by the high prevalence of malignancy and cardiovascular disease in our patients, which were the leading causes of long-term mortality subsequent to bacteremia in a study by Nielsen et al.[20]

AMR rates in the Bacthub database were similar to rates from French ECDC data for *K. pneumoniae, P. aeruginosa, S. aureus*, and *E. faecalis*.[1] Only rates of *P. aeruginosa* resistance to carbapenems and *E. coli* resistance to all tested antibiotics were clearly higher in the Bacthub database. Data from antibiotic resistance surveillance systems in France show that the Greater Paris area is among the regions with the highest rates of 3GC-R and ESBL production in *E. coli* isolates, in both community and hospital settings.[21,22] These results suggest a greater circulation of antibiotic-resistant *E. coli* strains in the Greater Paris area relative to the rest of France, which may partially explain our results. *P. aeruginosa* resistance to carbapenems is mainly due to chromosomal mutations, leading to loss or inactivation of the OprD porin and/or hyperexpression of efflux pumps.[23] The main risk factors for this resistance in the literature are exposure to antibiotics (especially carbapenems), medical devices, the presence of comorbidities, and ICU admission.[24] Therefore, our results likely reflect a higher incidence of HA bacteremia and greater exposure to healthcare and antibiotics in the Bacthub population. Indeed, the 14 hospitals included in our database are university-affiliated, with several being centers of expertise in infectious diseases. Thus, they treat particular and more comorbid patient populations than other health centers in the region. Despite these differences, overall trends in bacterial resistance over the study period in our database were consistent with both French and Europe-wide ECDC data.[1]

While most bacteremia databases are purely microbiological, the Bacthub database contains clinical data, allowing us to estimate bacteremia incidence at a multi-hospital level: 20–22 stays with bacteremia / 1000 stays. This estimate is higher than in previous multicenter surveys,[14,25] which could be explained by a progressive increase in bacteremia incidence over the years, which may in turn be linked to a progressive increase in life expectancy, as the elderly are most at risk.[18,25] Trends in incidence may further be explained by the particularities of the populations treated in the included hospitals.

Few studies have used CDWs to study acute infections, whether viral infections (SARS-CoV-2,[8] HIV,[9] HCV), bacteremia[7] or otherwise. Hospital data usually have numerous quality problems linked to the heterogeneity of reporting, extensive data missingness, incomplete data structuring, and the multiplicity of software programs used.[10] These factors often affect the validity of studies drawing upon hospital data. In our case, studying bacteremia using a highly representative microbiological database facilitated accurate infection reporting, as bacteremia diagnosis is reliable based purely on microbiological data. Moreover, we included a quality check to ensure that our results were representative of the hospitals included. After structuring our database, we described several major features of bacteremia that cannot typically be reported from microbiological databases alone, including patient and episode characteristics and clinical outcomes. The stability of our data over the years and the concordance of our main results with the literature suggest that this database can be reliably used for research purposes.

Our study has several limitations. Our definition of CO infection excluded patients discharged from any AP-HP hospital within 7 days prior to their admission. However, as AP-HP hospitals only cover one third of acute stays in the Île-de-France region, it is possible that some patients were wrongly classified as having CO infection. Furthermore, within CO bacteremia, we were not able to differentiate healthcare-associated from community-acquired bacteremia.[26] This may explain high rates of MRSA in the CO group, despite the low circulation of epidemic MRSA strains in the community in France over the study period. Finally, we used ICD-10 codes to determine the primary site of infection. No data are available to assess the reliability of this method, although widely used, in studies of medical-administrative databases.[3] With this method, one quarter of episodes were classified as primary bacteremia. This rate is higher than rates reported in two previous studies that used clinical and microbiological criteria to determine the primary site.[14,15] Nevertheless, the distribution of primary infection sites, as well as the distribution of bacterial isolates by infection site, are consistent with previous reports,[14] supporting the interpretability of diagnostic codes used in the database. An alternative approach could be to define primary infections based on a combination of clinical, biological, microbiological, and sometimes radiological criteria. However, the wide spectrum of potential primary sites in bacteremia, and the lack of reliable physiological data, prevented us from using this method. One way forward would be to extract this information from medical reports using natural language processing models, which to our knowledge has not yet been reported.

Overall, our experience with bacteremia shows that CDWs can be used as a reliable tool for conducting epidemiological research on acute bacterial infections, by combining diverse sources of information from clinical settings. Moreover, linking CDW data with various other databases, including diverse medical-administrative databases from both healthcare and community settings, could further increase their scope and research potential.[4,27,28] For instance, the Bacthub database could be enriched with data from the French National Health Data System, a medical-administrative database which records healthcare and office medicine reimbursements for the French population.[29] To overcome the main limitation of using CDWs for epidemiological research – the variable quality and completeness of hospital data – we call for increased collaboration between IT professionals and physicians, especially in the early stages of CDW development, to smooth the process of qualifying and structuring the data into a format suitable for research.

In conclusion, the Bacthub database is the first in France, and one of few on an international scale,[27,28] to combine highly accurate clinical and microbiological data for bacteremia episodes across a large multi-hospital group. The richness of the AP-HP CDW data allowed us to estimate in-hospital, 30-day and 90-day mortality rates for patients with community-onset and hospital-acquired episodes, and to provide a detailed overview of bacterial resistance data according to the bacteremia onset, the bacteria, and the primary infection site. The database offers several prospects for future research, including analyzing hospital care pathways of patients with bacteremia, and studying significant outcomes such as risk factors of mortality or recurrence of infection. In addition, these results could help guide the management of patients with bacteremia, including choice of antimicrobial therapy. Finally, linking the database to community-level data from the French National Health Data System, for instance, would facilitate description of care pathways and antibiotic exposure prior to hospitalization, allowing for more thorough study of community-acquired infections.

Abbreviations

3GC-R, third-generation cephalosporin resistant; AMR, antimicrobial resistance; AP-HP, Assistance Publique - Hôpitaux de Paris; AST, antibiotic susceptibility testing; CDW, Clinical data warehouses; CO, community-onset; DTI, digestive tract infection; ECDC, European Centre for Disease Prevention and Control; HA, hospital-acquired; ICD-10, International Classification of Diseases; ICU, intensive care unit; LRTI, lower respiratory tract infection; MRSA, methicillin-resistant S. aureus; STROBE, Strengthening the Reporting of Observational Studies in Epidemiology; UTI, urinary tract infection.

Data Sharing Statement

Data supporting this study can be made available on request from the AP-HP CDW, on the condition that the research project is accepted by the scientific and ethics committee of the AP-HP CDW.

Ethics Approval

This observational study uses routinely collected data. This study received approval from the Scientific and Ethical Committee of the Assistance Publique – Hôpitaux de Paris on March 28, 2019 and is registered on Clinicaltrials.gov (NCT04065750). The AP-HP clinical data warehouse initiative ensures patient information and informed consent regarding the different approved studies in accordance with European Regulation on data protection and authorization n°1980120 from the French Data Protection Authority (Commission Nationale de l'Informatique et des Libertés, CNIL).

Acknowledgments

The authors warmly thank the Health Data Hub staff, which provided essential administrative and technical support to the study, without which this database could not have been built. Our warm thanks also go to the AP-HP CDW staff (particularly the Innovation & Data (I&D) team) who implemented the CDW and greatly assisted us in selecting our population, building, and qualifying this database. We would like to also thank other AP-HP staff (particularly the Direction of Clinical research, Innovation and Relations with Universities and Research organizations (DRCI), the Clinical Research Unit of Paris Saclay Ouest (URCPSO), and all hospital departments that provided the data), and INSERM staff (particularly the Inserm Transfert, the Department of Legal Affairs (DAJ), and the Délégation Régionale Paris-IDF Sud). Finally, we thank David R.M. Smith for his critical review of the article and for proofreading the English.

Author Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

Funding

Doctoral fellowship was received by SA from the Ecole des Hautes Etudes en Santé Publique (EHESP, https://www.ehesp.fr/).

Disclosure

Funding received by SA had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results. LW received consulting fees from Pfizer, HEVA, IQVIA for unrelated projects. Other authors report no competing interests. The authors report no other conflicts of interest in this work.

References

1. Antimicrobial resistance in the EU/EEA (EARS-Net). Annual epidemiological report for 2019. Available from: https://www.ecdc.europa.eu/sites/default/files/documents/surveillance-antimicrobial-resistance-Europe-2019.pdf. Accessed December 8, 2022.
2. Murray CJ, Ikuta KS, Sharara F, et al. Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. *Lancet*. 2022;399(10325):629–655. doi:10.1016/S0140-6736(21)02724-0
3. Opatowski M, Tuppin P, Cosker K, et al. Hospitalisations with infections related to antimicrobial-resistant bacteria from the French nationwide hospital discharge database, 2016. *Epidemiol Infect*. 2019;147:e144. doi:10.1017/S0950268819000402
4. Wang XQ, Vincent BM, Wiitala WL, et al. Veterans Affairs patient database (VAPD 2014–2017): building nationwide granular data for clinical discovery. *BMC Med Res Methodol*. 2019;19(1):94. doi:10.1186/s12874-019-0740-x
5. Jannot AS, Zapletal E, Avillach P, Mamzer MF, Burgun A, Degoulet P. The Georges Pompidou University Hospital Clinical Data Warehouse: a 8-years follow-up experience. *Int J Med Inform*. 2017;102:21–28. doi:10.1016/j.ijmedinf.2017.02.006
6. Khalaf Hamoud A, Salah Hashim A, Akeel Awadh W. Clinical data warehouse a review. *IJCI*. 2018;44(2). doi:10.25195/2017/4424
7. McDanel JS, Roghmann MC, Perencevich EN, et al. Comparative effectiveness of cefazolin versus nafcillin or oxacillin for treatment of methicillin-susceptible staphylococcus aureus infections complicated by bacteremia: a nationwide cohort study. *Clin Infect Dis*. 2017;65(1):100–106. doi:10.1093/cid/cix287
8. Hoertel N, Sánchez-Rico M, Vernet R, et al. Observational study of haloperidol in hospitalized patients with COVID-19. *PLoS One*. 2021;16(2):e0247122. doi:10.1371/journal.pone.0247122
9. Kramer J, Hartman C, White D, et al. Validation of HIV -infected cohort identification using automated clinical data in the Department of Veterans Affairs. *HIV Med*. 2019:hiv.12757. doi:10.1111/hiv.12757

10. Daniel C, Serre P, Orlova N, Bréant S, Paris N, Griffon N. Initializing a hospital-wide data quality program. The AP-HP experience. *Comput Methods Programs Biomed*. 2019;181:104804. doi:10.1016/j.cmpb.2018.10.016
11. Quan H, Li B, Couris CM, et al. Updating and validating the Charlson comorbidity index and score for risk adjustment in hospital discharge abstracts using data from 6 countries. *Am J Epidemiol*. 2011;173(6):676–682. doi:10.1093/aje/kwq433
12. Annual Health Facilities Statistics (Statistique annuelle des établissements de santé). Available from: https://www.sae-diffusion.sante.gouv.fr/sae-diffusion/recherche.htm. Accessed December 8, 2022.
13. Weinstein MP, Towns ML, Quartey SM, et al. The clinical significance of positive blood cultures in the 1990s: a prospective comprehensive evaluation of the microbiology, epidemiology, and outcome of bacteremia and fungemia in adults. *Clin Infect Dis*. 1997;24(4):584–602. doi:10.1093/clind/24.4.584
14. Brun-Buisson C, Doyon F, Carlet J. Bacteremia and severe sepsis in adults: a multicenter prospective survey in ICUs and wards of 24 hospitals. French Bacteremia-Sepsis Study Group. *Am J Respir Crit Care Med*. 1996;154(3):617–624. doi:10.1164/ajrccm.154.3.8810595
15. Ortega M, Almela M, Martinez JA, et al. Epidemiology and outcome of primary community-acquired bacteremia in adult patients. *Eur J Clin Microbiol Infect Dis*. 2007;26(7):453–457. doi:10.1007/s10096-007-0304-6
16. Gradel KO, Jensen US, Schønheyder HC, et al.; For the Danish Collaborative Bacteraemia Network (DACOBAN). Impact of appropriate empirical antibiotic treatment on recurrence and mortality in patients with bacteraemia: a population-based cohort study. *BMC Infect Dis*. 2017;17(1):122. doi:10.1186/s12879-017-2233-z
17. Jensen US, Knudsen JD, Wehberg S, Gregson DB, Laupland KB. Risk factors for recurrence and death after bacteraemia: a population-based study. *Clin Microbiol Infect*. 2011;17(8):1148–1154. doi:10.1111/j.1469-0691.2011.03587.x
18. Sogaard M, Norgaard M, Dethlefsen C, Schonheyder HC. Temporal changes in the incidence and 30-day mortality associated with bacteremia in hospitalized patients from 1992 through 2006: a population-based cohort study. *Clin Infect Dis*. 2011;52(1):61–69. doi:10.1093/cid/ciq069
19. Nielsen SL, Lassen AT, Kolmos HJ, Jensen TG, Gradel KO, Pedersen C. The daily risk of bacteremia during hospitalization and associated 30-day mortality evaluated in relation to the traditional classification of bacteremia. *Am J Infect Control*. 2016;44(2):167–172. doi:10.1016/j.ajic.2015.09.011
20. Nielsen SL, Lassen AT, Gradel KO, et al. Bacteremia is associated with excess long-term mortality: a 12-year population-based cohort study. *J Infect*. 2015;70(2):111–126. doi:10.1016/j.jinf.2014.08.012
21. Primo mission 2019. Surveillance of bacterial resistance to antibiotics in community care and in institutions for the elderly dependent. Primo network: 2019 results. Saint-Maurice: Santé publique France; 2021. Available from: www.santepubliquefrance.fr. Accessed December 8, 2022.
22. Spares mission 2019. Surveillance of antibiotic consumption and bacterial resistance to antibiotics in healthcare institutions. Spares mission: 2019 results. Saint-Maurice: Santé publique France; 2021. Available from: www.santepubliquefrance.fr. Accessed December 8, 2022.
23. Botelho J, Grosso F, Peixe L. Antibiotic resistance in Pseudomonas aeruginosa – mechanisms, epidemiology and evolution. *Drug Resist Updates*. 2019;44:100640. doi:10.1016/j.drup.2019.07.002
24. Voor In 't Holt AF, Severin JA, Lesaffre EMEH, Vos MC. A systematic review and meta-analyses show that carbapenem use and medical devices are the leading risk factors for carbapenem-resistant Pseudomonas aeruginosa. *Antimicrob Agents Chemother*. 2014;58(5):2626–2637. doi:10.1128/AAC.01758-13
25. Nielsen SL. The incidence and prognosis of patients with bacteremia. *Dan Med J*. 2015;62(7):B5128.
26. Friedman ND. Health care–associated bloodstream infections in adults: a reason to change the accepted definition of community-acquired infections. *Ann Intern Med*. 2002;137(10):791. doi:10.7326/0003-4819-137-10-200211190-00007
27. Gradel K, Arpi M, Knudsen J, Schonheyder H, Ostergaard C, Sogaard M. The Danish Collaborative Bacteraemia Network (DACOBAN) database. *CLEP*. 2014;301. doi:10.2147/CLEP.S66998
28. Laupland KB, Schønheyder HC, Kennedy KJ, et al. Rationale for and protocol of a multi-national population-based bacteremia surveillance collaborative. *BMC Res Notes*. 2009;2(1):146. doi:10.1186/1756-0500-2-146
29. Tuppin P, Rudant J, Constantinou P, et al. Value of a national administrative database to guide public decisions: from the système national d'information interrégimes de l'Assurance Maladie (SNIIRAM) to the système national des données de santé (SNDS) in France. *Revue d'Épidémiologie et de Santé Publique*. 2017;65:S149–S167. doi:10.1016/j.respe.2017.05.004