Existing data sources for clinical epidemiology: The North Denmark Bacteremia Research Database

Henrik C Schønheyder1
Mette Søgaard1,2
1Department of Clinical Microbiology, Aalborg Hospital, Aarhus University Hospital, Aalborg, Denmark; 2Department of Clinical Epidemiology, Aarhus University Hospital, Aarhus, Denmark

Abstract: Bacteremia is associated with high morbidity and mortality. Improving prevention and treatment requires better knowledge of the disease and its prognosis. However, in order to study the entire spectrum of bacteremia patients, we need valid sources of information, prospective data collection, and complete follow-up. In North Denmark Region, all patients diagnosed with bacteremia have been registered in a population-based database since 1981. The information has been recorded prospectively since 1992 and the main variables are: the patient’s unique civil registration number, date of sampling the first positive blood culture, date of admission, clinical department, date of notification of growth, place of acquisition, focus of infection, microbiological species, antibiogram, and empirical antimicrobial treatment. During the time from 1981 to 2008, information on 22,556 cases of bacteremia has been recorded. The civil registration number makes it possible to link the database to other medical databases and thereby build large cohorts with detailed longitudinal data that include hospital histories since 1977, comorbidity data, and complete follow-up of survival. The database is suited for epidemiological research and, presently, approximately 60 studies have been published. Other Danish departments of clinical microbiology have recently started to record the same information and a population base of 2.3 million will be available for future studies.

Keywords: bloodstream infection, epidemiology, register, population-based

Introduction
Bacteremia is a clinical entity associated with detection of one or more micro-organisms in the blood, usually by culturing techniques. The entry of bacteria or fungi into the bloodstream can occur as a primary event that may lead to secondary infection of deep-sited organs and tissues or at a later stage if the pathogens are not contained at the site of infection by host defense mechanisms.1 Accordingly, the presence of viable micro-organisms in the circulating blood depends both on the capacity of micro-organisms to evade host defenses and on the competence of the patient’s immune system.

A number of cohort studies have confirmed the prognostic impact of bacteremia and fungemia,2,3 but the majority of the current studies are limited to one species or group of micro-organisms. Therefore a prospective recording of all patients with bacteremia and fungemia was started in North Jutland, Denmark in 1992 and in 1996 this became The North Denmark Bacteremia Research Database.

An historical perspective of bacteremia research
The term bacteremia was coined in the 19th century when bacteria and fungi were placed within the same taxonomic group.4 Since, generically, bacteremia includes...
both bacteria and fungi, we use the term accordingly in this review. Studies of bacteremia have a longstanding tradition. In particular, *Streptococcus pneumoniae*, $^7$ *Staphylococcus aureus*, $^8,9$ and enteric Gram-negative rods $^{10–12}$ have been studied extensively. In the 1970s and early 1980s, a number of thorough clinic-based outcome studies were conducted $^{5,13,14}$ and two of the most influential papers on bacteremia were published by Weinstein et al in 1983. $^{15,16}$ These twin papers included a detailed set of definitions that became a foundation for later research, including this database.

Information on bacteremia

Laboratory information systems can be used to identify patients with specific microbial agents or well defined groups thereof. However, when the entire spectrum of bacteremia is studied, the use of these data sources is more cumbersome, with the main challenge being distinction of the individual episodes. Information on bacteremia can also be retrieved from hospital discharge registries, but the codes vary between versions of the International Classification of Diseases and the sensitivity and specificity of the registration may be low. $^{17–19}$ Moreover, important microbiological information can be lost during the coding process. Combined microbiological and clinical databases dedicated to patients with bacteremia may overcome many of these problems. To our knowledge, the first databases that included all types of bacteremia were developed at St Thomas’s Hospital, London, England $^{20,21}$ and at Beilinson Hospital, Petah Tiqva, Israel $^{22,23}$ Both databases are clinic-based, however, and epidemiological studies can only be generalized with caution. In 1991, physicians at Aalborg Hospital in North Jutland, Denmark, used archived blood culture reports from the years 1981–1991 to establish a cohort of bacteremia patients. $^{24}$ This led to the development of a population-based bacteremia database with prospectively collected data. Although the initial aim was to improve the clinical service, $^{25}$ it was realized that the data collected presented a considerable research potential. $^{26}$

Setting and study population

“The of all the microbiological procedures performed in the laboratory, few are as important as the prompt recovery of micro-organisms from blood”. $^{27}$ This citation highlights the important clinical implications of blood cultures. $^{28}$ Still, the diagnosis requires teamwork between the physician who is in charge of the patient and the diagnostic laboratory that processes the blood cultures and provides the microbiological information. In Denmark, physicians in the departments of clinical microbiology take care of the notification of attending physicians, render advice on antibiotic therapy, and search for the focus of infection. This service makes the distinction between contaminants and significant blood culture isolates more straightforward. $^{13}$

Since 1992, the information obtained during these consultations has routinely been recorded in a structured format at the Department of Clinical Microbiology, Aalborg Hospital, which provides diagnostic microbiology to both general practitioners and public hospitals in the local health region (Table 1). The health region consists of mixed rural and urban areas with a modestly growing population. The Danish tax-paid health care system provides access, free of charge, to primary care by general practitioners as well as secondary and tertiary care in public hospitals. Currently, there are eight public hospitals in the region. Aalborg Hospital, an 800-bed university hospital, is the region’s referral hospital and serves as district hospital for the city of Aalborg. Patients in need of highly specialized treatment, including solid organ transplantation, are referred to one of three national centers outside the region.

Microbiological methods

Several improvements of the blood culture technique have been implemented over the years. $^{29–31}$ The most important being the introduction of a commercial broth culture system (ColorBact, Statens Serum Institut, Copenhagen, Denmark) in 1992. $^{30}$ In 1996 this system was replaced by the auto-

<table>
<thead>
<tr>
<th>Year</th>
<th>County of North Jutland</th>
<th>Number of bed days</th>
<th>Number of admissions</th>
<th>Number of blood cultures</th>
<th>Number of bacteremias</th>
</tr>
</thead>
<tbody>
<tr>
<td>1981</td>
<td>6173 km$^2$</td>
<td>481,501</td>
<td>746,110</td>
<td>3,327</td>
<td>368</td>
</tr>
<tr>
<td>1992</td>
<td>6173 km$^2$</td>
<td>485,787</td>
<td>550,916</td>
<td>8,916</td>
<td>671</td>
</tr>
<tr>
<td>2006</td>
<td>6173 km$^2$</td>
<td>495,090</td>
<td>483,646</td>
<td>17,047</td>
<td>1107</td>
</tr>
<tr>
<td>2007</td>
<td>7927 km$^2$</td>
<td>576,972</td>
<td>526,917</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Notes:**
1. North Jutland County became part of the larger North Denmark Region from January 1, 2007.
2. The years 2007–2009 encompassed a transitional period in which the clinical microbiological service did not include the region’s new hospitals. From 2010, the bacteremia database will again have full population-based coverage.
mated BacT/Alert® system (bioMérieux, Marcy l’Etoile, France). Each blood culture set includes two aerobic and one anaerobic bottle; for children up to 4 years of age, one pediatric bottle is used.

Positive blood cultures are subcultured and identified by a combination of conventional and commercial methods as previously described. Isolates likely to be contaminants according to the criteria set by Weinstein et al are submitted to a less extensive workup. Bacteria and fungi listed for national surveillance are referred to Statens Serum Institut, Copenhagen, together with unidentified micro-organisms. Pneumococcal serotypes, meningococcal serogroups/serotypes/serosubtypes, as well as S. aureus phage types (recently replaced by spa types), are obtained with courtesy of Statens Serum Institut. Molecular typing is rarely done, but isolates of the same species with a unique biochemical profile or antibiotic susceptibility pattern (antibiogram) are reported separately. Blood culture isolates, with the exception of contaminants, have been stored in glycerol broth at −80°C since 1995.

Data collection and variables
Definitions
Despite extensive literature on bacteremia, there was no coherent set of definitions on which a database could be based in 1992. We therefore developed a set of definitions, inspired by earlier clinic-based studies. We define the date on which the first positive blood culture is drawn as the start of bacteremia (T₀) and the baseline of follow-up. In patients with multiple positive blood cultures, we use one or more of the following criteria to distinguish a new episode: 1) a blood culture isolate that differs from the previous with regard to species and/or antibiogram, 2) a different focus of infection, or 3) an interval of at least 30 days between two positive blood cultures. We define polymicrobial bacteremia as an episode with two or more blood culture isolates of different species or antibiogram obtained within 48 hours. However, similar to Roberts, an exception to this is made for patients with a second isolate obtained within 25–48 hours if there is evidence of another focus of infection (eg, related to catheterization or surgery).

The focus is defined as the organ or tissue infected at the time when bacteremia becomes clinically apparent and a blood culture is drawn. Determining the focus is based on all microbiological and clinical evidence obtained during admission, including imaging studies. We do not distinguish between primary and secondary foci because this spread is rarely observed clinically.

The US Centers for Disease Control and Prevention (CDC) surveillance definitions of bacteremia pertain primarily to nosocomial infections and bacteremias are categorized as community-acquired if there is clinical evidence that the infection was present or incubating at the time of admission. We do not use a fixed time limit (eg, 48 or 72 hours) to distinguish between community-acquired and nosocomial bacteremias because such a criterion is not specified in the CDC guideline. Since 1992, we have considered patients with a hospital stay within 30 days prior to admission or who have regular hospital visits (eg, for hemodialysis or chemotherapy) as belonging to a separate ‘health care-associated’ group. This 30-day period is in agreement with Siegman-Igra et al. Others have chosen a period from 90 days to more than one year after a hospital stay in order to accommodate for long-term carriage of multiresistant pathogens such as methicillin-resistant S. aureus (MRSA). However, in our opinion, the categorization of health care-associated bacteremia should primarily account for a temporary impairment of the patient’s health related to the previous condition.

We consider antibiotic treatment given at the first notification as empirical and categorize it as appropriate if administered intravenously (with exception of fluoroquinolones and metronidazole) and if blood isolates are susceptible to one or more antibiotics given. If one or more isolates are resistant, or if the dose/route of administration is inadequate, treatment is categorized as inappropriate. An annually updated national Danish guideline serves as a reference for antibiotic dosing.

Informatics
From the outset, data were recorded in an SQL database (Paradox, Borland International, Scotts Valley, CA). Since 1996, North Jutland County has managed the database within the informatics framework of the health region. Transfer of data to a web server (Microsoft® Corp., Redmond, WA) has facilitated a partial integration with the microbiological laboratory information system (ADBakt and WWBakt, Autonik, Sweden) and the patient administrative system (PAS, Logica A/S, Denmark). Through the later system, the database is updated with information on vital status for residents in North Denmark Region and the neighboring Central Denmark Region. Since 2006, physicians have recorded the clinical information directly into the laboratory information system; eventually the bacteremia research database and the laboratory information system may become fully integrated.
Variables
Main variables recorded include: civil registration number; date of admission; clinical department; date and hour of venipuncture for the first positive blood culture (T₀); date and hour of the first notification of the attending physician; detection time in hours; a variable differentiating community-acquired, health care-associated, and nosocomial bacteremia (1992 onwards); focus of infection (from 1992); microbiological species; antibiogram; empirical and recommended antibiotic therapy; and appropriateness of antimicrobial treatment (from 1992). Table 2 provides a complete list of the fields recorded in the database for each bacteremic episode.

Time period
All patients diagnosed with bacteremia since January 1, 1981 are registered in the database; data covering 1981 to 1991 was collected retrospectively from archived blood culture reports, but since 1992 the information has been recorded concurrently with the clinical episode. Between 1981 and 2008, 22,556 cases have been recorded (5114 [23%] from 1981–1991, 15,062 [67%] from 1992–2006, and 2380 [11%] from 2007–2008). Figure 1 shows a modest increase in annual numbers from 1981 to 1991, followed by a steep rise up to 2004. The time from 2007 to 2009 represents a transitional period for which we do not have full population-based coverage in the new North Denmark Region.

Methodological aspects
Strengths and weaknesses
The database covers all patients diagnosed with bacteremia in North Denmark Region and that it incorporates a complete register-based follow-up for mortality is one of the major strengths. Use of the civil registration number assigned to all Danes makes it possible to link bacteremia data to other medical and socioeconomic Danish registries including the Danish National Registry of Patients (with data on all hospital admissions and discharge diagnoses since 1977 and all outpatient visits since 1995), the Danish Civil Registration System (which contains data on immigration, emigration, and death), the Prescription Database for North Denmark Region (which retains key information on prescriptions for reimbursable drugs dispensed from pharmacies), and a clinical biochemistry database (LABKA, CSC Scandihealth, Denmark, which holds data recorded since 1997 for patients both in primary and hospital care). This makes it possible to build large cohorts with detailed longitudinal data that include full hospitalization history, comorbidity data, and complete long-term follow-up data. The population base makes the database especially suited for matched cohort studies of community-acquired bacteremias. The Danish Civil Registration System facilitates the selection of controls from the underlying population using incidence density sampling.47,48

Table 2 List of the fields recorded in the database for each bacteremic episode

<table>
<thead>
<tr>
<th>Unique bacteremia episode number</th>
</tr>
</thead>
<tbody>
<tr>
<td>The patient’s 10 digit personal civil registry number</td>
</tr>
<tr>
<td>Age</td>
</tr>
<tr>
<td>Sex</td>
</tr>
<tr>
<td>Clinical department (at time of venipuncture)</td>
</tr>
<tr>
<td>Clinical speciality</td>
</tr>
<tr>
<td>Date of admission</td>
</tr>
<tr>
<td>Transfer (including from nursing home)</td>
</tr>
<tr>
<td>Laboratory identification number of sentinel blood culture</td>
</tr>
<tr>
<td>Time of the venipuncture (date and hour)</td>
</tr>
<tr>
<td>Time of first notification (date and hour)</td>
</tr>
<tr>
<td>Number of positive blood culture bottles</td>
</tr>
<tr>
<td>Number of blood culture bottles inoculated</td>
</tr>
<tr>
<td>Focus of infection</td>
</tr>
<tr>
<td>Evidence to support the focus: clinical (including imaging) (Y/N)</td>
</tr>
<tr>
<td>Evidence to support the focus: microbiological (Y/N)</td>
</tr>
<tr>
<td>Evidence to support the focus: pathoanatomic (Y/N)</td>
</tr>
<tr>
<td>Microbiological specimens other than blood (Y/N)</td>
</tr>
<tr>
<td>Place of acquisition</td>
</tr>
<tr>
<td>Date of death</td>
</tr>
<tr>
<td>Alive day 30 (Y/N)</td>
</tr>
<tr>
<td>Alive day 180 (Y/N)</td>
</tr>
</tbody>
</table>

Accessory clinical information:
- Chronic urinary catheter
- Intravenous line
- Permanent intravenous catheter
- Chronic urinary catheter
- Surgical procedure (Y/N)
- Trauma patient (Y/N)

Comorbid diseases (Y/N)
- Solid cancer
- Hematological cancer
- Liver cirrhosis
- Diabetes mellitus
- Renal failure
- Dialysis

Cumulative number of bacteremic episodes (1, 2, 3, ...)
Number of blood culture isolates

Relation: Antibiotic therapy
- Antibiotic(s) given at first notification
- Antibiotic(s) recommended at first notification
- Appropriateness of empirical antibiotic therapy
- Antibiotic(s) given at second notification
- Antibiotic(s) recommended at second notification

Blood culture isolate number: #
Species
- Type (serotype, serogroup, phage type, spa type, etc)

Relation: Antibiotic susceptibility (SIR) of blood culture isolate number: #
- Antibiotic (up to 24 antibiotics)

Abbreviation: SIR, susceptible-indeterminate-resistant.
Furthermore, we consider both the concurrent recording of information and the strict definitions to be advantageous. The definition most important for the number and pattern of bacteremias is probably the duration of the bacteremic episode. In the literature, the duration varies from 2 to more than 30 days\textsuperscript{15,49–52} and although a 30-day limit is arbitrary, it should allow time for treatment of the underlying focus and clinical recovery. Moreover, positive blood cultures obtained within this period are likely due to either ineffective antibiotic therapy or an unresolved focus of infection and do rarely constitute an entirely new clinical episode.

Finally, efforts are continuously made to validate the bacteremia information recorded in the database and there is very little missing data.

The main weakness of the database is the lack of clinical data on acute disease severity including SIRS criteria,\textsuperscript{53} as well as other clinical scores. We hope that implementation of electronic medical records will allow more detailed characterization of patients with bacteremia and their treatments for future studies.

**Research from the database**

The predominant research themes have been the risk and outcome of bacteremia. In the study of community-acquired bacteremia, we were among the first to use the Charlson Index to adjust for comorbidity.\textsuperscript{31} Quite early, we also became interested in the long-term prognosis of bacteremia.\textsuperscript{54,55}

A study from 1999 described the secular trends in incidence and mortality of bacteremia in North Jutland County during 1981 to 1994.\textsuperscript{30} During that period the incidence increased by 100\% (from 76 to 153 per 100,000 person-years), whereas the 30-day case fatality rate remained stable (23.6\%, 95\% confidence interval [CI]: 22.6–24.6).

Another priority has been to address bacteremia in patients with prevalent diseases such as liver cirrhosis,\textsuperscript{64} alcohol abuse,\textsuperscript{57} diabetes mellitus,\textsuperscript{47,58} and hematological malignancies.\textsuperscript{59,60} Thomsen et al\textsuperscript{47} found that patients with diabetes had a substantially increased risk for bacteremia caused by enteric Gram-negative rods (adjusted odds ratio [OR] 2.9, 95\% CI: 2.4–3.4). Among patients with bacteremia, diabetes was also associated with a poorer prognosis (adjusted mortality rate ratio [MRR] 1.3, 95\% CI: 0.9–1.8).

Recently, we used the bacteremia database to describe the impact of age and comorbidity on short-term mortality from community-acquired bacteremia.\textsuperscript{61} We found older age and greater comorbidity to be associated with higher 30-day mortality. Thirty-day mortality rates increased from 11\% in patients younger than 65 years to 21\% in patients aged 80 years and older. Likewise, the 30-day mortality rate
in patients with low comorbidity was 7% but increased to 23% in patients with high comorbidity. We further found that increasing age-related comorbidity did not explain the effect of age. This prognostic effect of age on short-term bacteremia mortality has also been demonstrated in patients with hematological malignancies.

The database has further been used to study trends in antibiotic resistance and antibiotic treatment of bacteremia. Pedersen et al. showed that antibiotic treatment within 90 days prior to admission was associated with antibiotic-resistant blood culture isolates in patients with community-acquired bacteremia; this observation is important because the impact of antibiotic treatment on resistance is rarely studied in patients with systemic community-acquired infections. From 1999 to 2003, an international working group used data both from the Danish and the aforementioned Israeli bacteremia database to develop a medical decision support system for antibiotic treatment.

Access for other researchers
The database is public and researchers with permission from the Danish Data Protection Agency will be granted access to the data. Moreover, we would be interested in collaborating with external research institutions under the provisions of the Data Protection Agency. Data files are kept for research purposes by the Department of Clinical Epidemiology, Aarhus University, while the original data files are kept by the North Denmark Region. Potential collaborators are invited to contact the corresponding author.

Perspectives
Knowledge of the occurrence and outcome of bacteremia is important to establish its burden of illness to guide appropriate allocation of limited health care and research resources. Population-based studies are optimal to describe the epidemiology of bacteremia because the well defined catchment area allows determination of incidence rates and risk factors. However, our knowledge about preventable or modifiable risk factors and prognostic factors for bacteremia in a population-based setting remains limited. We hope that in the future, joint efforts by researchers in disciplines including clinical epidemiology, molecular epidemiology and pharmaco-epidemiology, clinical microbiology, infectious diseases, and clinical medicine will provide new insights into the etiology, course, and prevention of bacteremia. Future work may also uncover potential sources of infection and reservoirs in humans, animals, and nature.

The continuous development of the North Denmark Bacteremia Research Database is a priority and includes full integration with departmental and hospital information systems. The database has inspired interregional cooperation on bacteremia research and key information on bacteremia cases is currently recorded in the same format by three Danish departments of clinical microbiology on a routine basis. This holds promise of a population base of approximately 2.3 million inhabitants for future bacteremia studies.

Other regions worldwide have reported on population-based bacteremia data including Australia, Canada, Iceland, Finland, and the United States. Recently, an international collaboration was established that represents six regions in four countries on three continents with a combined annual surveillance population of more than 8 million residents. This collaboration will facilitate comparison of interregional differences in incidence, risk factors, outcomes, and resistance rates. Moreover, the large surveillance population allows the study of rare isolates and assists early detection of emerging pathogens. Finally, the recognition of the importance of global trade and international travel in the spread of multiresistant organisms has highlighted that infectious disease control is a global concern.

Acknowledgments
We are indebted to many collaborators and our special thanks are due to Henrik Toft Sørensen, Lars Pedersen, Tove Højbjerg, Gitte Pedersen, Brian Kristensen, Reimar W Thomsen, Mette Nørgaard, Anne Marie Thulstrup, Kim Oren Gradel, Jenny Dahl Knudsen, and Christian Østergaard.

Inger Lindhardt (†) and Lena Mortensen have assisted meticulously with keeping the bacteremia database update.

We thank the medical directors for Aalborg Hospital, Per Østergaard Jensen and Tove Nilsson, for their unfailing support. The support from the Health Informatics Department in both North Jutland County and North Denmark Region is also gratefully acknowledged. Finally, we acknowledge the financial support from the North Jutland County, North Denmark Region, and Det Obelske Familiefond.

Disclosure
The Bacteremia Research Database is supported by the health region (both North Jutland County and North Denmark Region) and Department of Clinical Epidemiology, Aarhus University. In recent years, Aalborg Hospital has provided funding for a part-time assistant for keeping the database
updated. Additional funding has been provided by private foundations.

References


68. Jensen US, Østergaard C, Schønheyder HC, Knudsen JD. Creating a collaborative network for the study of bacteremia in Denmark: frequency of recurrence with the same and different micro-organisms. European Congress of Clinical Microbiology and Infectious Diseases, Helsinki; 2009 May 16–19 [Poster 1437].
