Supplementary Online Content


eAppendix. Detailed methods.

This supplementary material has been provided by the authors to give readers additional information about their work.
eAppendix

DETAILED METHODS

Data Source

We conducted a retrospective cohort study using the Nationwide Inpatient Sample (NIS), an all-payer administrative dataset produced by the Agency for Healthcare Research and Quality’s Healthcare Cost and Utilization Project. Through the use of sampling methods (weighting and stratification based on geographic region and facility type), the NIS is designed to be representative of short-stay hospitalizations at non-Federal facilities in the United States. In 1999, 24 states participated; by 2008 the number had increased to 42. Each year of the NIS dataset includes discharges from approximately 1,000 hospitals and contains approximately 8 million records (20% of short-stay, non-Federal hospitalizations). As the largest all-payer administrative inpatient database in the United States, the NIS is well suited for the study of rare conditions such as IE. NIS data have previously been used to evaluate the prevalence and temporal trends of Staphylococcus aureus infections. The dataset includes patient demographics, diagnosis and procedure codes, length of stay, discharge disposition, and total inpatient facility charges. Beginning in 2002, NIS datasets also include comorbidity data obtained from a validated diagnosis code-based algorithm. We used NIS data over a ten-year period, from the first year in which species-specific organism codes were available for staphylococcal infections (1999) to the most current data available (2008). The sampling methodology of NIS was unchanged during the study period, allowing for consistent comparisons of trends in bacterial IE incidence and epidemiology.

Cohort selection, organism identification, and comorbidity determination

Hospital stays related to bacterial IE were identified by the presence of International Classification of Diseases, Ninth Edition, Clinical Modification codes 421.0, 421.1, 421.9 or 996.61.
on any discharge diagnosis field, combining two previously employed strategies.\textsuperscript{3,6,7} For those stays identified as IE-related, the etiological agent of infection was determined by the presence of organism-specific infection codes (e.g. 041.x) or organism-specific bacteremia codes (038.x). IE-related admissions with multiple organism codes were considered polymicrobial. Streptococcal and enterococcal organisms share a common bacteremia code. This table details the coding strategy, developed with an experienced inpatient coder.

**Coding Criteria**

<table>
<thead>
<tr>
<th>Discharge Diagnosis / Procedure Code</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Infective Endocarditis</strong></td>
</tr>
<tr>
<td>421.0, 421.1, 421.9, 996.61</td>
</tr>
</tbody>
</table>

**Organism Criteria**

- **Staphylococcus aureus**
  - 038.11, 038.12, 041.11, 041.12
- Unspecified Staphylococci
  - 038.10, 041.10
- Non-\textit{S. aureus} Staphylococci
  - 038.19, 041.19
- Streptococcus / Enterococcus
  - 038.0, 041.00-041.09
- Other
  - 041.3-041.9, 038.42, 003.1, 038.43, 116.0, 083.0, 002.0

**Statistical Analysis**

The NIS consists of data from hospital discharge records. Bacterial IE incidence was estimated using the rate of IE-related \textit{discharges} per 100,000 US population-years instead of the traditional \textit{incident cases} per 100,000 population-year metric, as was previously done for studies of infection prevalence using administrative data.\textsuperscript{7} In order to limit the over-counting of cases caused by patients being transferred from one facility to another, we excluded hospitalizations where the discharge disposition was listed as another inpatient facility. Rates were calculated for each...
calendar quarter, based on date of patient discharge, and the population denominator was adjusted annually using data from the United States Census Bureau.\textsuperscript{8, 9} Trends in admission rate were evaluated using joinpoint regression models, which allow for changes in the trend over time.\textsuperscript{10} We used the natural log of admission rate as the dependent variable in these analyses, so trends were estimated in terms of the quarterly average percent change.

For bacterial IE admissions in which an etiologic organism could be identified from the data, weighted regression models were used to compare baseline patient characteristics based on organism. To evaluate the impact of bacterial etiology on in-hospital mortality, we used logistic regression models to adjust for patient age, gender, primary payer, and clinical comorbidities. Since comorbidity data were included in NIS from 2002 onward, these models excluded 1999-2001 admissions and excluded records in which other covariate data were missing. Because odds ratios produced by logistic regression are difficult to interpret when the dependent variable (in this case, mortality) is common\textsuperscript{11} regression results are presented in terms of average marginal effects, which estimates the average effect of bacterial etiology in terms of the incremental difference in probability of death. All analyses were conducted treating streptococcal / enterococcal IE as the reference group.

A two-sided alpha level of 0.05 was considered significant for hypothesis testing. The dataset was constructed in SAS System, version 9.22 (SAS Institute). Statistical analyses were performed in Stata/IC, version 11.2 (Statacorp) and Joinpoint Regression Program, version 3.4.3 (National Cancer Institute). This study was ruled exempt from review by the University of North Carolina at Chapel Hill Biomedical Institutional Review Board.
References


