Control-Group Selection Importance in Studies of Antimicrobial Resistance: Examples Applied to Pseudomonas aeruginosa, Enterococci, and Escherichia coli

Anthony D. Harris, Matthew H. Samore, Marc Lipsitch, Keith S. Kaye, Eli Perencevich, and Yehuda Carmeli

1Department of Epidemiology and Preventive Medicine, University of Maryland School of Medicine, and 2VA Maryland Health Care System, Baltimore; 3Department of Internal Medicine, University of Utah, Salt Lake City; 4Harvard School of Public Health, Boston, Massachusetts; 5Duke University Medical Center, Durham, North Carolina; and 6Tel Aviv Sourasky Medical Center, Israel

We aimed to illustrate the importance of control-group selection on the results of risk factor analysis for (1) imipenem-resistant Pseudomonas aeruginosa, (2) vancomycin-resistant enterococci (VRE), and (3) ampicillin-sulbactam–resistant Escherichia coli. Case patients were compared with 2 different control groups: patients with the susceptible form of the organism (type 1), and control patients among whom the case patients arose during the same period as the case patients (type 2). Comparison of case patients who had imipenem-resistant P. aeruginosa with type-1 control patients identified use of imipenem (odds ratio [OR], 27.1) and quinolones (OR, 3.25) as a risk factor for selection of antimicrobial resistance, and comparison of the same case patients with type-2 control patients identified imipenem (OR, 6.34). When case patients with VRE were compared with type-1 and with type-2 control patients, use of vancomycin was identified as a risk factor (OR, 4.38 and 2.77, respectively). Comparison of case patients who had ampicillin-sulbactam–resistant E. coli compared with type-1 control patients identified ampicillin-sulbactam (OR, 2.71) and quinolones (OR, 2.72), and comparison with type-2 control patients identified ampicillin-sulbactam (OR, 1.68). The selection of control patients from the potentially suboptimal control type 1 can falsely identify certain antibiotics and overestimate the OR of the resistance-defining antibiotic.

The spread of antimicrobial resistance is a major threat to public health. The case-control study design is often used to identify risk factors for the acquisition of antibiotic-resistant organisms. In 3 recent publications [1–3], we have highlighted the theoretical importance of control-group selection in studies of risk factors for antibiotic resistance.

In studies that have analyzed the use of individual antimicrobial agents as individual risk factors for the isolation of antimicrobial-resistant organisms in the hospital, 2 types of control groups are chosen most frequently: patients with cultures positive for the antibiotic-susceptible form of the organism of interest (type 1), and a sampling of patients admitted to the hospital who are at risk of acquiring the antimicrobial-resistant organism (type 2). In our other articles, we pointed out that the significance of the ORs may not be meaningfully inferred from studies that rely on control groups that consist of patients with the antimicrobial-susceptible form of the organism, because the effect...
of antibiotics may be distorted by the use of control patients who have the antimicrobial-susceptible form of the organism. The selection bias introduced by the use of control patients with antimicrobial-susceptible organisms is likely to have the largest impact on antibiotics that are active against susceptible but not against the resistant form of the organism [2].

In this article, we illustrate and quantify the importance of control-group selection on the results of risk-factor analyses for 3 resistant organisms: (1) imipenem-resistant *Pseudomonas aeruginosa*, (2) vancomycin-resistant enterococci (VRE), and (3) ampicillin-sulbactam–resistant *Escherichia coli*. We are particularly interested in how the magnitude of the risk associated with use of specific antibiotics may change and how use of certain antibiotics may or may not be identified as a risk factor when different control groups are used. We were particularly interested in assessing the magnitude of the change in risk for the antibiotics imipenem (for study 1), vancomycin (for study 2), and ampicillin-sulbactam (for study 3) because of their activity against the susceptible organisms associated with control patients and lack of activity against the resistant organisms associated with case patients.

**MATERIALS AND METHODS**

**Study design.** Two case-control studies were performed for each organism of interest, involving 2 groups of patients: (1) unique patients (a patient could be a case patient only once) who had clinical culture results that were positive for the resistant organism were compared with unique patients who had clinical culture results that were positive for the antimicrobial-susceptible organism, and (2) unique patients who had clinical culture results that were positive for the resistant organism were compared with randomly selected control patients. The 3 organisms studied were imipenem-resistant *P. aeruginosa*, VRE, and ampicillin-sulbactam–resistant *E. coli*.

Thus, in study 1a, case patients who had clinical culture results positive for imipenem-resistant *P. aeruginosa* were compared with control patients who had clinical culture results positive for imipenem-resistant *P. aeruginosa*. In study 1b, case patients who had clinical culture results positive for imipenem-resistant *P. aeruginosa* were compared with randomly selected control patients. In study 2a, case patients who had clinical culture results positive for VRE were compared with control patients who had clinical culture results positive for vancomycin-susceptible enterococci. In study 2b, case patients who had clinical culture results positive for VRE were compared with randomly selected control patients. In study 3a, case patients who had clinical cultures positive for ampicillin-sulbactam–resistant *E. coli* were compared with control patients who had clinical culture results positive for ampicillin-sulbactam–susceptible *E. coli*. In study 3b, case patients who had clinical culture results positive for ampicillin-sulbactam–resistant *E. coli* were compared with randomly selected control patients.

**Case definition.** Studies 1 and 2 were performed at the University of Maryland Medical System, a 609-bed tertiary care teaching hospital in Baltimore, Maryland. Study 3 was performed at the Beth Israel Deaconess Medical Center West Campus (BIDMC), a 320-bed tertiary care teaching hospital in Boston, Massachusetts. The studies were approved by the institutional review boards at each hospital.

Case patients for each of the 3 studies were defined as patients for whom there was nosocomial isolation of a resistant organism from clinical cultures—for study 1, imipenem-resistant *P. aeruginosa*; for study 2, VRE; and for study 3, ampicillin-sulbactam–resistant *E. coli*.

For studies done at the University of Maryland, patients were admitted and discharged during the period of 1 January 1998 through 1 July 2000. For the studies done at the BIDMC, patients were admitted and discharged during the period of October 1993 through September 1997. In all studies at both institutions, patients from whom resistant isolates were obtained within the first 48 h after admission were excluded. Isolates obtained from surveillance cultures were also excluded.

**Control-group definition.** In study 1a, control patients were defined as patients from whom imipenem-susceptible *P. aeruginosa* were isolated. In study 2a, control patients were defined as patients from whom vancomycin-susceptible enterococci were isolated. In study 3a, control patients were defined as patients from whom ampicillin-sulbactam–susceptible *E. coli* were isolated. Control patients with positive results of clinical cultures in the first 48 h after admission were excluded.

In study 1b, control patients were defined as patients for whom clinical cultures did not yield imipenem-resistant *P. aeruginosa* during their hospital stay. In study 2b, control patients did not have VRE isolated. In study 3b, control patients did not have ampicillin-sulbactam–resistant *E. coli* isolated.

For studies 1b and 2b, control patients were selected from the same medical or surgical services in which case patients resided when the case patients had resistant organisms isolated. Six control patients from the same service were chosen for each case patient’s service location. Control patients were admitted during the same period as the case patients. For study 3b, control patients were a computer-generated random selection of 5% of all patients who were admitted during the same period as case patients. Patients hospitalized for <48 h were excluded from the control groups of studies 1b, 2b, and 3b.

**Risk factors investigated.** Data for studies 1 and 2 were collected from administrative, pharmacy, and laboratory computerized databases by means of a relational database management system. The relational database at the University of Maryland is maintained by the Information Technology Group of
the University of Maryland. The pharmacy, microbiology, and medical demographics tables in the relational database have been validated against the medical records for >400 patients admitted during the period of October 1997 through January 2001. In addition, data for a 10% sample of case patients and a 5% sample of control patients in the present study were validated with medical charts. The positive and negative predictive values of the data are >99%. The details of the database for study 3 are outlined in a study published elsewhere [4]. In brief, a central data repository is maintained by the Information Technology Group of the BIDMC and consists of administrative, pharmacy, and laboratory computer data [5].

Variables studied included age, sex, underlying diseases or comorbid conditions (determined according to the coding of the International Classification of Diseases, Ninth Revision), transfer from another hospital, intensive care unit stay before the outcome of interest, surgery before the outcome of interest, time at risk before the outcome of interest, and treatment with antimicrobial drugs. The definitions for time at risk were as follows: (1) for case patients, the length of hospital stay before the isolation of a resistant organism; (2) for control patients in studies 1a, 2a, and 3a, the length of hospital stay before the isolation of an antimicrobial-susceptible organism; and (3) for control patients in studies 1b, 2b, and 3b, complete length of hospital stay.

For the case patients in studies 1 and 2, treatment with antimicrobial drugs was included in the analysis only when the antimicrobial drugs were given ≤7 days before isolation of the resistant organism. For the control patients in studies 1a and 2a, treatment with antimicrobial drugs was included in the analysis only when the antimicrobial drugs were given ≤7 days before isolation of the antimicrobial-susceptible organism. For control patients in studies 1b and 2b, treatment with antimicrobial drugs during the 7 days before discharge from the hospital was included. The rationale behind the choice of 7 days was to avoid including antibiotics that patients had received during the initial phase of a lengthy hospitalization in the analysis. However, because the 7-day period was arbitrarily chosen, statistical analyses were repeated using a 14-day time period, and the results obtained were very similar; thus, for concision, only data for the 7-day-period are presented in the Results section. The antimicrobial drugs for which data were statistically analyzed were piperacillin-tazobactam; ampicillin-sulbactam; first-, second-, and third-generation cephalosporins; macrolides; aminoglycosides; and quinolones.

In study 3, for case patients, treatment with antimicrobial drugs was analyzed as a potential risk factor only when the drugs were given between the time of admission and of isolation of the resistant organisms. For the control patients in study 3a, treatment with antimicrobial drugs was included in the analysis only when the antimicrobial drugs were given between the time of admission and of isolation of the antimicrobial-susceptible organism. For control patients in study 3b, treatment with any antimicrobial drug throughout the entire hospital stay was included. The antimicrobial drugs that were statistically analyzed were ampicillin, ampicillin-sulbactam, piperacillin, piperacillin-tazobactam, cefazolin, cefuroxime and cefotetan, ceftriaxone and ceftazadime, aminoglycosides, quinolones, and imipenem.

**Statistical analysis.** All statistical analyses were performed by use of SAS software, version 7 (SAS Institute). Bivariate analyses were performed separately for each of the variables. ORs and 95% CIs were calculated for binomial variables; P values were calculated by Fisher’s exact test, for binomial variables; c2 test, for categorical variables with >2 subgroups; and Student’s t test or the Wilcoxon rank-sum test, for continuous variables. Variables with a P value of <.1 in the bivariate analysis were included in a logistic-regression model for multivariable analysis. A forward-selection process was used. Risk factors were checked for confounding and colinearity. Confounding variables were included in the multivariable models if covariate inclusion changed the coefficient of any statistically significant variable in the logistic-regression model by ≥10%. All tests were 2-tailed, and a P value of ≤.05 was considered significant in the multivariable model.

**RESULTS**

In study 1a, there were 120 case patients with clinical culture results positive for imipenem-resistant *P. aeruginosa* and 662 control patients with clinical culture results positive for imipenem-susceptible *P. aeruginosa*. Study 1b included the same 120 case patients and 770 randomly selected control patients. In study 2a, there were 406 case patients with clinical culture results positive for VRE and 1586 control patients with clinical culture results positive for vancomycin-susceptible *enterococci*. Study 2b included the same 406 case patients and 3134 randomly selected control patients. In study 3a, there were 175 case patients with clinical culture results positive for ampicillin-sulbactam–resistant *E. coli* and 577 control patients with clinical culture results positive for ampicillin-sulbactam–susceptible *E. coli*. Study 3b included the same 175 case patients and 934 randomly selected control patients.

For studies 1a and 1b, 51 case patients (43%) had received imipenem before the isolation of imipenem-resistant *P. aeruginosa*. For study 1a, 20 (3%) of the control patients had received imipenem before the isolation of imipenem-susceptible *P. aeruginosa*. For study 1b, 43 randomly selected control patients (6%) had received imipenem. For study 1a, the bivariate OR for imipenem was 27.12 (P < .0001). For study 1b, the bivariate OR for imipenem was 12.50 (P < .0001).

For studies 2a and 2b, 154 case patients (38%) had received vancomycin before the isolation of VRE. For study 2a, 127
Table 1. Multivariable analyses of risk factors for the isolation of imipenem-resistant *Pseudomonas aeruginosa*.

<table>
<thead>
<tr>
<th>Control group, risk factor</th>
<th>OR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1a</td>
<td></td>
</tr>
<tr>
<td>Use of imipenem</td>
<td>27.12 (13.91–52.90)</td>
</tr>
<tr>
<td>Use of aminoglycosides</td>
<td>2.38 (1.40–4.05)</td>
</tr>
<tr>
<td>Use of quinolones</td>
<td>3.25 (1.92–5.49)</td>
</tr>
<tr>
<td>Surgery</td>
<td>0.42 (0.21–0.85)</td>
</tr>
<tr>
<td>1b</td>
<td></td>
</tr>
<tr>
<td>Use of imipenem</td>
<td>6.34 (3.66–11.00)</td>
</tr>
<tr>
<td>Use of aminoglycosides</td>
<td>3.28 (1.98–5.42)</td>
</tr>
<tr>
<td>Time at risk, days</td>
<td>1.03 (1.01–1.04)</td>
</tr>
<tr>
<td>Intensive care unit stay</td>
<td>3.85 (2.16–6.86)</td>
</tr>
</tbody>
</table>

a Patients with imipenem-susceptible *P. aeruginosa*.
b Randomly selected patients.

(8%) of the control patients had received vancomycin before the isolation of vancomycin-susceptible enterococci. For study 2b, 313 (10%) of the randomly selected control patients had received vancomycin. For study 2a, the bivariate OR for vancomycin was 7.54 (P < .0001). For study 2b, the bivariate OR for vancomycin was 5.43 (P < .0001).

For studies 3a and 3b, 31 case patients (18%) had received ampicillin-sulbactam before the isolation of ampicillin-sulbactam–resistant *E. coli*. For study 3a, 43 control patients (7%) had received ampicillin-sulbactam before the isolation of ampicillin-susceptible *E. coli*. For study 3b, 99 (11%) of the randomly selected control patients had received ampicillin-sulbactam. For study 3a, the bivariate OR for ampicillin-sulbactam was 2.67 (P < .0001). For study 3b, the bivariate OR for ampicillin-sulbactam was 1.82 (P = .007).

The results of the multivariable risk-factor analyses for imipenem-resistant *P. aeruginosa* (studies 1a and 1b) are presented in table 1. Study 1a, which had a susceptible-organism control group, identified use of the following antibiotics as a risk factor: imipenem (OR, 27.12; 95% CI, 13.91–52.90), aminoglycosides (OR, 2.38; 95% CI, 1.40–4.05), and quinolones (OR, 3.25; 95% CI, 1.92–5.49). Study 1b, which included randomly selected control patients, identified use of the following antibiotics as a risk factor: imipenem (OR, 27.12; 95% CI, 13.91–52.90), aminoglycosides (OR, 2.38; 95% CI, 1.40–4.05), and quinolones (OR, 3.25; 95% CI, 1.92–5.49). Study 1b, which included randomly selected control patients, identified use of the following antibiotics as a risk factor: imipenem (OR, 27.12; 95% CI, 13.91–52.90), aminoglycosides (OR, 2.38; 95% CI, 1.40–4.05), and quinolones (OR, 3.25; 95% CI, 1.92–5.49). Study 1b, which included randomly selected control patients, identified use of the following antibiotics as a risk factor: imipenem (OR, 27.12; 95% CI, 13.91–52.90), aminoglycosides (OR, 2.38; 95% CI, 1.40–4.05), and quinolones (OR, 3.25; 95% CI, 1.92–5.49).

The results of the multivariable risk-factor analyses for VRE (studies 2a and 2b) are presented in table 2. In study 2a, which had control patients with vancomycin-susceptible enterococci, the OR for vancomycin use was 4.38 and the 95% CI was 3.24–5.93, whereas, in study 2b, which included randomly selected control patients, the OR was 2.86, and the 95% CI was 2.19–3.73. The results for the other antibiotics are given in table 2.

The results of the multivariable risk factor analyses for ampicillin-sulbactam–resistant *E. coli* (studies 3a and 3b) are presented in table 3. Study 3a, which had a susceptible-organism control group, identified use of the following antibiotics as a risk factor: ampicillin-sulbactam (OR, 2.71; 95% CI, 1.64–4.46) and quinolones (OR, 2.72; 95% CI, 1.16–6.37). Study 3b, which had a randomly selected control group, identified use of the following antibiotics as a risk factor: ampicillin-sulbactam (OR, 1.68; 95% CI, 1.02–2.77) and ampicillin (OR, 2.69; 95% CI, 1.08–6.69), and first-generation cephalosporins were determined to be protective (OR, 0.31; 95% CI, 0.18–0.53).

**DISCUSSION**

The importance of control-group selection in case-control studies has been outlined in a number of epidemiological publications since 1985 [1, 6–9]. However, only recently has the importance of control-group selection been reexplored as it pertains to studies of antimicrobial resistance [1–3, 5, 10]. These latter articles outlined the potential selection bias that arises if patients with antimicrobial-susceptible organisms are used as control patients. Despite these principles, the majority of case-control studies choose patients with the susceptible organism for their control group [2].

Our aim in the present study was to demonstrate, by use of concrete examples, the selection bias that may arise by the choice of a suboptimal control group for studies that address...
the research question, “What are the risk factors for acquiring antibiotic-resistant pathogen X among hospitalized patients?”

In studies 1, 2, and 3, we demonstrated that the selection bias has the largest effect on antibiotics that have activity against susceptible (but not resistant) organisms—that is, imipenem for imipenem-resistant *P. aeruginosa*, vancomycin for VRE, and ampicillin-sulbactam for ampicillin-sulbactam–resistant *E. coli*.

For example, in both the bivariate and multivariate analyses of the study of imipenem-resistant *P. aeruginosa*, the OR for imipenem was much larger in study 1a than it was in study 1b. We hypothesize that the OR was much larger in study 1a because imipenem protects patients against having subsequent culture results positive for imipenem-susceptible *P. aeruginosa* and, thus, prevents patients from becoming members of the susceptible-organism control group.

Similarly, for study 2, which analyzed the risk factors for acquiring VRE, there was a statistically significant difference in the OR for vancomycin use in the randomly selected control group. This may be because the receipt of vancomycin prevents patients from having subsequent positive culture results for vancomycin-susceptible enterococci. A similar explanation explains the differences in ampicillin-sulbactam ORs in studies 1a and 1b.

In addition to affecting the magnitude of risk, the selection of the susceptible-organism control group may also falsely identify risk factors. In study 1a, use of quinolones was identified as a risk factor for the development of imipenem-resistant *P. aeruginosa*, whereas, in study 1b, it was not. In this study population, there was evidence that imipenem-resistant *P. aeruginosa* was more likely to be resistant to ciprofloxacin than was imipenem-susceptible *P. aeruginosa* (49% vs. 16% of patient isolates). Thus, ciprofloxacin treatment may have made individuals less likely to have imipenem-susceptible *P. aeruginosa* relative to imipenem-resistant *P. aeruginosa*. In study 3a, use of quinolones was identified as a risk factor, whereas, in study 3b, it was not. In this study population, there was evidence that ampicillin-sulbactam–resistant *E. coli* was more likely to be resistant to fluoroquinolones than was ampicillin-sulbactam–susceptible *E. coli* (3.2% vs. 1.0% of patient isolates). Therefore, fluoroquinolone treatment may have made individuals less likely to have ampicillin-sulbactam–susceptible *E. coli* relative to ampicillin-sulbactam–resistant *E. coli* [5].

The conclusion that use of a control group that consists only of subjects colonized or infected with the susceptible form of the organism introduces bias and is therefore not recommended has certain qualifications. For example, when the hypothesis focuses on the emergence of drug resistance, the susceptible-organism group becomes the source population. Thus, to answer the question, “What are risk factors for emergence of antibiotic-resistance in pathogen X among patients previously infected or colonized with antibiotic-susceptible pathogen X,” the appropriate and optimal case-control study design is to select, as control patients, individuals with the susceptible organism and no subsequent resistant organism and, as case patients, individuals with the resistant organism who previously had the susceptible form of the organism [2].

Another caveat is that, for some pathogens, the OR derived from the comparison of patients with susceptible organisms and patients with resistant organisms may be the more appropriate measurement of the impact of antimicrobial use at the population level [3].

An additional limitation is that, because the control patients in these studies were not screened with active surveillance cultures for the presence of the antibiotic-resistant organism, it is possible that some of these patients might actually have been case patients. This type of misclassification would make case and control patients more similar by including case patients in the control group and could lead to an underestimation of the strength of associations.

Length of hospital stay is an important confounding variable, a point that has been outlined in an article elsewhere [2]. In the context of a case-control study, controlling for length of hospital stay may be accomplished either by including it as a variable in a logistic-regression model or by matching for time at risk during the process of control patient selection. The latter method is referred to as “risk-set sampling.” Whether these alternative methods of adjusting for length of hospital stay yield disparate results awaits further study. An alternate design, namely the case-cohort method, could also be used to analyze risk factors, and the statistical method of Kaplan-Meier analysis could be used to adjust for length of hospital stay or time at risk.

A limitation of the present study was that we were unable

---

**Table 3. Multivariable analyses of risk factors for the isolation of ampicillin-sulbactam–resistant *Escherichia coli*.**

<table>
<thead>
<tr>
<th>Control group, risk factor</th>
<th>OR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3a* Use of ampicillin-sulbactam</td>
<td>2.71 (1.64–4.46)</td>
</tr>
<tr>
<td>3a Use of quinolones</td>
<td>2.72 (1.16–6.37)</td>
</tr>
<tr>
<td>3b Randomly selected patients. Age</td>
<td>1.01 (1.00–1.03)</td>
</tr>
<tr>
<td>3b Use of ampicillin</td>
<td>2.69 (1.08–6.69)</td>
</tr>
<tr>
<td>3b Use of ampicillin-sulbactam</td>
<td>1.68 (1.02–2.77)</td>
</tr>
<tr>
<td>3b Use of first-generation cephalosporins</td>
<td>0.31 (0.18–0.53)</td>
</tr>
<tr>
<td>3b Male sex</td>
<td>0.33 (0.23–0.48)</td>
</tr>
<tr>
<td>3b Hepatic disease</td>
<td>1.89 (1.08–3.22)</td>
</tr>
<tr>
<td>3b Intensive care unit stay</td>
<td>2.42 (1.62–3.63)</td>
</tr>
<tr>
<td>3b Surgery</td>
<td>2.07 (1.36–3.17)</td>
</tr>
<tr>
<td>3b Transfer from a different hospital</td>
<td>1.41 (0.96–2.07)</td>
</tr>
</tbody>
</table>

* Patients with ampicillin-sulbactam-susceptible *E. coli*.

* Randomly selected patients.
to assess the role of patient-to-patient transmission. Horizontal transfer and colonization pressure have been demonstrated to be important in a number of studies [11]. However, it is likely that not accounting for patient-to-patient transmission would bias toward the null—that is, among patients who acquire the organism from another patient, the importance of antibiotics as causal components may be diminished, because patient-to-patient transmission is the principal causal factor.

In the present study, we have demonstrated that the choice of control group affects the identification of use of antimicrobial agents as risk factors and the magnitude of the effects in case-control studies of antibiotic resistance. Selection of control groups from the appropriate population at risk and not solely enrollment of patients with cultures positive for susceptible organisms should provide a more accurate assessment of the magnitude of risk associated with specific antimicrobial agents by minimizing selection bias. We believe that improvements and refinements in the methodology of case-control studies of antibiotic resistance will lead to the identification of valid risk factors and appropriate targets for interventions aimed at curbing the increasing emergence of antibiotic resistance.

Acknowledgment

We thank Colleen Reilly for database maintenance and data extraction.

References