Investigation of sources of potential bias in laboratory surveillance for anti-microbial resistance

Kevin B. Laupland MD MSc1,3,5
Terry Ross5
Johann D.D. Pitout MB ChB2,4
Deirdre L. Church MD PhD1,2,4
Daniel B. Gregson MD1,2,4

Departments of Medicine1, Pathology and Laboratory Medicine2, and Community Health Sciences3, University of Calgary and Calgary Health Region, Division of Microbiology4, Calgary Laboratory Services, Centre for Anti-microbial Resistance5, University of Calgary, Calgary Health Region, and Calgary Laboratory Services, Calgary, Alberta, Canada.

Manuscript submitted 23rd March, 2007
Manuscript accepted 27th May, 2007


Abstract

Purpose: There are a number of biases that may influence the validity of laboratory-based surveillance for antimicrobial resistance. Our objective was to evaluate the potential magnitude of bias in reporting of etiologic agents and their resistance rates associated with inclusion of multiple patient samples and non-random timing and location of sampling.

Methods: All urine cultures submitted to a regional laboratory in the Calgary Health Region during 2004 and 2005 were studied. Comparisons were then made using either the overall cohort or different subgroups compared with the “reference” or gold standard population where only the first isolate per patient per year per species was included.

Results: Overall 56,897 organisms were cultured at ≥10^4 cfu/mL from 53,548 samples from 35,890 patients; 39,835 organisms were included in the reference cohort. Escherichia coli was reported in 37,246 (65.5%) of overall cohort and 28,257 (70.9%) of the reference cohort. Therefore, the overall cohort resulted in a relative underestimation of the importance of E. coli as the principal cause of urinary tract infections by 8%. Similarly, reported rates of resistance to antimicrobial agents most notably ciprofloxacin [6,480/52,544 (12.3%) vs. 2,647/37,086 (7.1%)], gentamicin [2,991/48,070 (6.2%) vs. 1,567/34,608 (4.5%)], and ceftri-
gence of resistance, respectively. There are a number of potential biases that may affect the results of laboratory-based surveillance systems. They include but are not limited to failure to remove multiple repeated or duplicate isolates, incorrect or inconsistent laboratory technique, inadequate sampling size, and selection bias related to the timing or location of sampling.

While there are several potential biases that may be important, limited data exist on whether they have a practical influence on results or conclusions. The objective of this study was to investigate the magnitude of selected potential sources of bias related to surveillance methodology that may influence validity of laboratory-based anti-microbial resistance data. To achieve this goal we compared an overall urinary tract culture cohort and selected subgroups with reference “gold standard” laboratory-based surveillance cohort where only one sample per patient per year was included.

Methods

Study population

The Calgary Health Region is a health system that provides all medically necessary, publicly funded health care to the more than one million residents of the cities of Calgary and Airdrie and multiple nearby small towns, villages, and hamlets in a total area of 37,000 square km. All significantly positive (≥10^4 colony forming units/mL) urinary tract cultures identified by Calgary Laboratory Services during 2004 and 2005 were included in the overall cohort; cultures with lesser counts were excluded. Calgary Laboratory Services performs nearly all (>95%) standard microbiology testing from hospitals, nursing homes, physicians’ offices, and community collection sites in the Calgary Health Region. The “gold standard” or reference cohort was drawn from the overall cohort but was limited to include only the first isolate per patient per year per species. Because this study involved an analysis of laboratory data with externally recognizable individual patient identifiers removed, specific institutional ethics review was not sought.

Laboratory procedures and definitions

Urine was cultured using standard clinical laboratory methodology. Identification to the species level and antimicrobial susceptibility testing was performed routinely using Vitek (Vitek AMS; bioMérieux Vitek Systems, Inc., Hazelwood, MO) and interpreted using the Clinical and Laboratory Standards Institute criteria for broth dilution. Isolates with intermediate susceptibility were considered resistant. For the purposes of this study, coagulase negative staphylococci (other than Staphylococcus saprophyticus), Lactobacillus species, and cultures that contained more than two different species were considered contaminants and were excluded from analysis. All positive cultures were identified using a data extraction from the laboratory database (Cerner Pathnet Classic version 306, Kansas City, MO).

Basic demographic data including age, sex and the sending site was obtained on all patients. Patients were then classified to sending location as outpatients (clinics, physician’s offices, and outpatient collection sites), nursing homes, and hospitals (emergency departments and hospital inpatients). For hospitalized patients, the date of admission and discharge were obtained. Hospital onset urinary tract infections were those that occurred in patients with first culture positive more than 2 days after hospital admission. Community onset urinary tract infections were all others and included outpatients, nursing home residents, emergency department patients, and those admitted to hospital with first positive culture within the first two days of admission. Given that this study was designed to assess laboratory-based surveillance, detailed clinical information on patients was not obtained.

Analysis

Data were exported from the laboratory database and managed in Access 2003 (Microsoft Corp., Redmond, WA). All analyses were conducted using Stata 9.0 (Stata Corp., College Station, TX). To assess the effect of assessment of multiple samples from patients, a priori defined subgroups were developed where only the first sample per 1, 2, 4, and 6 months were included and were then compared descriptively with the
reference cohort that only included the first species isolate per patient per year and the overall culture cohorts. To assess potential selection bias related to sampling during different time periods and locations, differences in distribution of organisms and resistance to antibiotics were compared among subgroups (testing time of day, days of week, season and outpatient, nursing home, and hospital setting) of the reference cohort. In these comparisons, differences in proportions were compared using the $\chi^2$ test with $P<0.05$ deemed to represent statistical significance.

**Results**

**Overall cohort**

During the two years of the study, 155,550 patients submitted 290,411 urine specimens for culture: 56,897 organisms were cultured at $10^4$ colony forming units/mL in 53,548 samples from 35,890 patients. The majority (30,565; 85%) of the patients were female and the median age was 59.9 [inter-quartile range (IQR); 32.2-85.7] yr. Sixty-six percent (37,824) of the cultures were obtained from outpatients, 25% (14,416) from hospital based patients, and 8% (4,657) were from nursing home residents. Nearly one-half (6,789; 47%) of the hospital cultures were obtained from emergency department patients and of the 7,287 admitted patients, 5,439 (75%) were classified as hospital-onset.

**Effect of period of repeated sample submission**

A modest difference was observed with the proportion of species causing urinary tract infections based on the period of repeated sample submission testing. The main difference between the overall cohort (37,246/56,897; 65.5%) and the gold standard cohort (28,257/39,835; 70.9%) was in the proportion of urinary tract infections attributed to *E. coli*. Therefore, the overall cohort resulted in a relative underestimation of the importance of *E. coli* as the principal cause of urinary tract infections by 8%.

Other than gentamicin (1.3 for females vs. 1.5 fold for males) there was no difference observed in relation to gender. When only *E. coli* isolates were analyzed, the same higher rates of resistance were observed with the exceptions that ceftriaxone was 1.3 and nitrofurantoin was 1.8 fold higher compared to the gold standard cohort.
toin was 1.7 fold increased among all isolates compared with the reference cohort.

### Evaluation of the timing of sampling

There were differences in the species distribution depending on whether cultures were completed during different periods of the day, days of the week, and quarters of the year \((P<0.001)\). The proportion of cultures from ambulatory, hospital, and nursing home sites reported during 0000h to 0800h was 71.8%, 24.8%, and 3.4%, during 0801h to 1600h was 70.8%, 24.1%, and 5.2%, and during 1601h to 2359h was 58.2%, 32.7%, and 9.2%, respectively. The proportion of the predominant pathogen *E. coli* in specimens completed 0000h to 0800h was 4,181/4,921 (85.0%), 0801h to 1600h was 23,977 (69.1%), and 1601h to 2359h 99/196 (50.5%; \(P<0.001\)). In the reference cohort, 70.8% of cultures were from ambulatory patients, 24.2% were from hospital patients, and 5.0% were submitted from nursing home residents. On Monday and Tuesday a higher proportion of cultures were from hospital patients (39.7%) and a lower proportion were from ambulatory patients (56.0%). The proportion of *E. coli* was lowest early in the week [Monday and Tuesday 4,181/4,921 (85.0%)], reached a peak mid-week [Wednesday and Thursday 9,463/12,989 (72.9%)], and then decreased again [Friday to Sunday 14,228/19,901 (71.5%); \(P<0.01\)]. There was a small, lower proportion of *E. coli* isolated in the first quarter (January to March: 6,985/10,032; 69.6%) than in the other three quarters of the year (April to December: 21,272/29,803; 71.4%; \(P=0.001\)).

A large relationship between the timing during the day of completion of a culture and resistance rates was observed \((P<0.001\) for all anti-microbials) with rates as much as ten fold higher in the evening as shown in Table 2. Similarly there was a relationship between day of the week and reporting of resistance rates \((P<0.001\) for each individual anti-microbial). This was largely due to increased rates of resistance in specimens reported on Monday/Tuesday compared with other days of the week (Table 2). Although relationships between resistance rate and quarters of the year were observed for ciprofloxacin, TMP/SMX, and nitrofurantoin \((P<0.001\) for each), the differences in magnitude were very small.

### Influence of location of sampling

The proportional distribution of species was related to whether samples were obtained from outpatients, hospitals, or nursing homes \((P<0.001)\). The predominant pathogen *E. coli* was isolated in 76.5% (21,589/28,212), 59.1% (5,697/9,637), and 48.9% (971/1,986) of cases, respectively. The proportion of *E. coli* isolated was 76.5% (21,589/28,212), 59.1% (5,697/9,637), and 48.9% (971/1,986) of cases, respectively. The differences in magnitude were very small.

| **TABLE 2. Effect of timing of sampling on the reported rate of resistance to common anti-microbials used to treat urinary tract infection, Calgary Health Region, 2004-2005** |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| **Agent**       | **Reference cohort** | **0000-0800** | **0801-1600** | **1601-2359** | **Mon/Tues†** | **Weds/Sun†** |
| Ciprofloxacin   | 2,646/37,086 (7.1) | 159 (3.3)      | 2,402 (7.5)   | 85 (49.4)     | 618 (9.7%)     | 2,028 (6.6%)   |
| TMP/SMX        | 4,916/34,201 (14.4) | 582 (12.3)     | 4,278 (14.6) | 56 (42.8)     | 900 (15.9%)     | 4016 (14.1%)   |
| Cefazolin      | 3,276/34,132 (9.6) | 253 (5.4)      | 2,967 (10.1) | 56 (43.8)     | 710 (12.6%)     | 2566 (9.0%)    |
| Gentamicin     | 1,567/36,608 (4.5) | 67 (1.4)       | 1,468 (4.9)  | 32 (23.5)     | 368 (6.7%)      | 1,181 (4.1%)   |
| Nitrofurantoin | 2,972/36,635 (8.1) | 274 (5.8)      | 2,678 (8.4)  | 20 (12.3)     | 532 (8.5%)      | 2,440 (8.0%)   |
| Ceftriaxone    | 889/32,745 (2.7) | 25 (0.5)       | 819 (2.9)    | 45 (38.5)     | 240 (4.5%)      | 649 (2.4%)     |

* \(\chi^2 P<0.001\) for each anti-microbial resistance rate comparison and time period of day
† \(\chi^2 P<0.001\) for each comparison except nitrofurantoin \(P=0.2\)
among major hospital site among those admitted to one of the major acute care centers in the Calgary Health Region \((P<0.001)\) as shown in Table 3.

**Discussion**

In this study we document the magnitude of a number of potential biases of laboratory-based surveillance studies of resistance arising from inclusion of multiple specimens from patients and differential timing and location of sampling. These data emphasize the need for rigor in the design of surveillance systems methodology, and raise concerns surrounding the interpretation of results from surveillance studies that do not protect against these and other forms of bias. Reporting of biased data may result in either under or over-estimated rates of resistance among organisms that may have the important effect of falsely influencing clinicians to use inadequate or excessively broad spectrum antibiotic therapy, respectively.

Our data largely confirm the work of other investigators that have identified considerable bias with failure to exclude duplicate or multiple specimens with urinary tract infections\(^9\), *Staphylococcus aureus* surveillance\(^10,14\), Gram-negative bacilli\(^15\), and *Klebsiella pneumoniae* and *Acinetobacter baumannii* isolates from inpatients\(^16\). The Clinical and Laboratory Standards Institute recommends that with laboratory data where the purpose is in recommending empiric therapy, only the first isolate per species per patient per analysis period be included \(^6\). However, the timing of removal of duplicate isolates for urinary tract infections remains to be defined. In one study of female enrollees of a health maintenance organization in Seattle, Gupta and colleagues arbitrarily excluded repeat isolates within one month and found that nearly all isolates using this duplicate exclusion criteria were from clinically relevant new episodes\(^17\). It should be recognized that this was a clinical study and that laboratory-based surveillance systems, in particular the large multinational programs that provide the vast majority of antimicrobial resistance data, do not have access to clinical information to define clinically discrete episodes. From the perspective of both the distribution of species causing infection as well as with anti-microbial susceptibilities, our data suggest that a one-year duplicate removal interval be used for laboratory-based studies. Evidence in support of this is that we observed increased frequencies compared with the reference cohort even with a six-month elimination period (Table 1). While further study is

### Table 3. Effect of location of sampling on the reported rate of resistance to common anti-microbials used to treat urinary tract infections, Calgary Health Region, 2004-2005

<table>
<thead>
<tr>
<th>Agent</th>
<th>Out-patients(^a)</th>
<th>Nursing homes(^a)</th>
<th>Hospitalized(^a)</th>
<th>Hospital A(^†)</th>
<th>Hospital B(^†)</th>
<th>Hospital C(^†)</th>
<th>Hospital D(^†)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ciprofloxacin</td>
<td>1,390 (5.3%)</td>
<td>338 (18.7%)</td>
<td>918 (10.4%)</td>
<td>28/882 (3.2%)</td>
<td>363/3286 (11.1%)</td>
<td>213/1,876 (11.4%)</td>
<td>286/2,285 (12.5%)</td>
</tr>
<tr>
<td>TMP/SMX</td>
<td>3,557 (14.2%)</td>
<td>248 (15.8%)</td>
<td>1,111 (14.6%)</td>
<td>148/805 (18.4%)</td>
<td>374/2,733 (13.7%)</td>
<td>258/1,654 (15.6%)</td>
<td>258/1,941 (13.3%)</td>
</tr>
<tr>
<td>Cefazolin</td>
<td>1,947 (7.8%)</td>
<td>259 (16.5%)</td>
<td>1,070 (14.1%)</td>
<td>82/802 (10.2%)</td>
<td>437/2,723 (16.1%)</td>
<td>237/1,650 (14.4%)</td>
<td>272/1,934 (14.1%)</td>
</tr>
<tr>
<td>Gentamicin</td>
<td>994 (4.0%)</td>
<td>135 (8.3%)</td>
<td>438 (5.6%)</td>
<td>36/817 (4.4%)</td>
<td>177/2,840 (6.2%)</td>
<td>100/1,688 (5.9%)</td>
<td>103/2,007 (5.1%)</td>
</tr>
<tr>
<td>Nitrofurantoin</td>
<td>1,733 (6.6%)</td>
<td>380 (21.6%)</td>
<td>859 (10.0%)</td>
<td>49/867 (5.7%)</td>
<td>368/3,168 (11.6%)</td>
<td>169/1,836 (9.2%)</td>
<td>249/2,212 (11.3%)</td>
</tr>
<tr>
<td>Ceftriaxone</td>
<td>487 (2.0%)</td>
<td>53 (3.7%)</td>
<td>349 (4.9%)</td>
<td>19/775 (2.5%)</td>
<td>155/2,524 (6.1%)</td>
<td>73/1,554 (4.7%)</td>
<td>92/1,807 (5.1%)</td>
</tr>
</tbody>
</table>

\(^a\) \(\chi^2 P<0.001\) for comparison among outpatients, nursing homes, and hospitalized patients for each anti-microbial except \(P=0.2\) for TMP/SMX.

\(^†\) \(\chi^2 P<0.002\) for comparison among hospitals for each anti-microbial except \(P=0.14\) for gentamicin
required to better define the optimal procedures for duplicate/multiple isolate removal, this is an important potential bias in surveillance systems that must be considered in the assessment of the validity of such data.

There has been little attention paid to investigating potential bias related to the timing and location of collection of surveillance data. We found that the time of reporting of cultures may have a major influence on the species distribution and isolate susceptibilities in surveillance studies, with much lower rates of *E. coli* and higher rates of resistance reported later in the day and evening. It is important to emphasize that the objective of this study was to identify sources of bias in surveillance. We did not collect data to explain why these biases may occur. It is likely that the higher rate of resistance and lower rates of *E. coli* isolation in the evening and early days of the week were due to higher proportions of hospital isolates assessed during those times. In addition, we also speculate that organisms that are difficult to speciate or with higher resistance rates may be reported later in the day after they are reviewed by the medical microbiologist. Nonetheless, it is important to recognize the major potential bias associated with the timing and location of sampling because multi-center surveillance studies typically request that participating centers submit some predefined number of consecutive samples to the coordinating study center; a non-biased random sample is not specified. Our data suggest that, depending on the time of day and day of the week (and to a much lesser extent season of year), the consecutive samples are drawn would have potential to lead to results that are an invalid representation of the overall species distribution and resistance profile of that laboratory.

In addition to the timing of collection of samples from a participating surveillance site, the selection of the site(s) of sample is also an important determinant on results. It is neither surprising nor a new finding that organism distribution and rates of resistance differ substantially among outpatients, hospitalized patients, and nursing home residents. Surveillance systems usually will define the participating laboratories. However, it should be recognized that results obtained from hospital-based laboratories that have variable associated community based or nursing home catchments may provide disparate and non-representative results. Furthermore, if community onset infections are under surveillance, it is important to study isolates from both community based labs as well as those from patients with community onset disease assessed at hospitals. Based on multicentred surveillance studies conducted in hospitals, extended-spectrum β-lactamase producing organisms were widely believed to be hospital onset pathogens. However, once studies that included both hospital and community-based laboratories were conducted, it was evident that these organisms were major community onset pathogens.

We observed considerable differences among isolates submitted from the four major acute care hospitals in our study. This is an important consideration in multicentred studies because academic, university-affiliated, tertiary care hospital laboratories or those with a particular problem with resistant organisms may be more likely to participate than smaller community hospital laboratories. Ideally, if multicentred or national studies are to be conducted, all or randomly sampled isolates from all laboratories within the surveillance catchment area should be included. However, such population-based studies are not usually feasible. An alternative, methodologically rigorous, but rarely employed approach would be to select randomly a sample of laboratories from among all those in the surveillance catchment area. The potential for study laboratory selection or “participant” bias must be recognized and results from such studies interpreted with care.

There are important methodological aspects of this study that merit discussion. First, it is a strength that our overall cohort included all urinary cultures (≥10⁴ colony forming units/mL) obtained from patients in our region. Because we did not perform sampling from within this population, selection bias was minimized. However, it should be recognized that our determination of species distribution and resistance rates in urinary tract infections in our population was dependent on a sample being submitted to the laboratory. Many low risk patients diagnosed clinically with
uncomplicated cystitis were likely treated empirically without cultures being obtained and samples were more likely to have been sent from patients with complicated disease. Thus our results likely overestimate the true rates of resistance in all urinary tract infections. This is a limitation inherent in all laboratory-based studies\textsuperscript{17}. A second consideration with our observations surrounding multiple sampling and selection bias due to selected timing and location bias is that it is possible that the magnitude of the biases we observed may be lesser or greater with other types of cultures or in other jurisdictions. Since this data is novel, the onus is on further studies in our and other centers to challenge our findings.

In conclusion, using a large laboratory-based population cohort of urinary tract isolates we found considerable biases may result in reporting of surveillance data if repeat specimens are included and random sampling is not performed. While further confirmation of the magnitude of these biases in other patient populations with different infection types is needed, these data argue that care must be taken in interpreting results of surveillance studies that do not protect from these and other potential sources of bias.

**Acknowledgments**

No external funding was obtained for this study.

**References**

beta-lactamase production in nosocomial infections. 

19. Pitout JD, Gregson DB, Church DL, Elsayed S,  
Laupland KB. Community-wide outbreaks of clonally  
related CTX-M-14 beta-lactamase-producing Escheri-  
chia coli strains in the Calgary health region. *J Clin  

20. Pitout JD, Nordmann P, Laupland KB, Poirel L.  
Emergence of Enterobacteriaceae producing extended-  
spectrum beta-lactamases (ESBLs) in the community.  

Correspondence to:

Dr. Kevin B. Laupland  
Room 719, North Tower, Foothills Medical Centre  
1403 – 29th Street NW  
Calgary, Alberta, CANADA T2N 2T9  
email: kevin.laupland@calgaryhealthregion.ca