Antimicrobial resistance surveillance systems: Are potential biases taken into account?

Olivia Rempel MLT BHSc, Johann DD Pitout MBChB M Med(Micro Path) FFPath(SA), Kevin B Laupland MD MSc FRCP(C)


BACKGROUND: The validity of surveillance systems has rarely been a topic of investigation.

OBJECTIVE: To assess potential biases that may influence the validity of contemporary antimicrobial-resistant (AMR) pathogen surveillance systems.

METHODS: In 2008, reports of laboratory-based AMR surveillance systems were identified by searching Medline. Surveillance systems were appraised for six different types of bias. Scores were assigned as ‘2’ (good), ‘1’ (fair) and ‘0’ (poor) for each bias.

RESULTS: A total of 22 surveillance systems were included. All studies used appropriate denominator data and case definitions (score of 2). Most (n=18) studies adequately protected against case ascertainment bias (score = 2), with three studies and one study scoring 1 and 0, respectively. Only four studies were deemed to be free of significant sampling bias (score = 2), with 17 studies classified as fair, and one as poor. Eight studies had explicitly removed duplicates (score = 2). Seven studies removed duplicates, but lacked adequate definitions (score = 1). Seven studies did not report duplicate removal (score = 0). Eighteen of the studies were considered to have good laboratory methodology, three had some concerns (score = 1), and one was considered to be poor (score = 0).

CONCLUSION: Contemporary AMR surveillance systems commonly have methodological limitations with respect to sampling and multiple counting and, to a lesser degree, case ascertainment and laboratory practices. The potential for bias should be considered in the interpretation of surveillance data.

Key Words: Antimicrobial resistance; Bias; Epidemiology; Incidence; Risk

The emergence of organisms with resistance to antimicrobial agents is a paramount contemporary health care issue, and surveillance has been recognized as a fundamental component in their control (1-3). Surveillance information enables the assessment of the burden of disease, determination of risk factors, and identification of temporal trends in occurrence and resistance patterns of infectious diseases. Such information may be used to aid policy-makers in their decisions regarding health services and research funding allocation, to guide efforts into means of prevention and control, and to direct emergency antimicrobial therapy recommendations. However, the value of surveillance information is predicated on its reliability and validity. Invalid surveillance data risks wasting health care resources through misguided efforts, and may result in patient harm through inappropriate use of antimicrobial agents. Although surveillance data have been widely published and used by researchers and decision makers, little attention has been devoted to the assessment of their validity (3-10).

We previously explored six potential biases that may influence surveillance systems including bias related to the use of inadequate or inappropriate denominator data (I); case definitions (II) and case ascertainment (III); sampling bias (IV); failure to deal with multiple occurrences (V); and biases related to laboratory practice and procedures (VI) (3). To our knowledge, there have been no previous systematic attempts to examine contemporary antimicrobial-resistant (AMR) organism surveillance systems with regard to how well they protect from bias. The objective of the present study was to assess current surveillance systems for AMR organisms against the six main biases identified in our previous literature review (3).

METHODS

A sample of current surveillance systems was obtained by searching the Medline database through the PubMed interface using the terms “antimicrobial resistance” and “surveillance”. The search was limited to
include only human-based surveillance articles written in English and published in 2008. After identifying all studies meeting the above criteria, abstracts were initially screened by one author (OR) under the supervision of another (KBL) to identify relevant studies for further review. Given that surveillance is generally expected to have an ongoing component, one-time resistance surveys were excluded (11), and only studies that included ongoing, multicentric, laboratory-based surveillance that reported on at least one AMR organism were included. When multiple studies from the same surveillance system were obtained, only the first study appearing in the search results for 2008 was included for analysis. However, if methods were referenced elsewhere, then these reports were retrieved and reviewed.

A scoring guide was developed to assess the surveillance studies according to the six biases identified in our previous literature review (3). Scores were assigned as 0, 1 or 2 for each of the six potential biases. A score of 0 was assigned where measures to protect from bias were either poor or not reported. Studies that reported some measures to protect against the bias under consideration were scored as 1. Study methodologies that were well-protected against bias were scored as 2. After assigning scores for each of the six biases, the scores were summed to obtain a total final score from zero to 12.

Two reviewers (OR and KBL) independently reviewed and scored the selected studies. Scores were based on the selected publications alone; supplemental searching for added detail through other means, such as the Internet or by contacting authors, was not performed. If there were potentially overlapping areas of bias recognized, then each area was considered separately such that studies could not lose points for the same issue more than once. For example, if an issue surrounding case definitions also directly led to problems with case ascertainment, then a reduced score was recorded for the case definition, but case ascertainment was scored assuming an adequate case definition. Once the two independent reviews were completed, discrepancies were resolved through consensus with a third reviewer (JDDP).

Analysis was primarily descriptive. The weighted kappa statistic was calculated to assess the level of agreement between the two independent reviewers—both overall and for each of the six biases examined (12).

### RESULTS

Initially, 459 abstracts were screened and, of these, 22 fulfilled the study inclusion criteria and were reviewed in detail and scored. Of these, there were five studies from the United States, four studies from Europe, three studies from Asia, three studies from Australia and New Zealand, three studies from Canada, two studies from multiple continents, and one study from each of Africa and South America. The included studies and their consensus scores are summarized in Table 1. The overall median score was 10 (range 7 to 11), and the weighted kappa score among the two reviewers was 0.82.

### Denominator data and case definitions

The weighted kappa for each of denominator data and case definitions was 1. All studies scored ‘2’ for denominator data and case definition. Of these, 18 studies examined resistance among all isolates obtained (13-30). The other four were population based, and the denominator included all patients at risk for antimicrobial resistance (31-34). Because all studies based their case definition on resistance using well-established guidelines, case definition was not an issue for any of the surveillance studies examined.
Case ascertainment

For case ascertainment, the weighted kappa statistic was 0.60. Eighteen of the studies scored ‘2’ for case ascertainment (13-19,21,23-25,27-33). These studies were either population based or reported on the proportion of isolates that were resistant. It was deemed that all cases meeting the case definition were included, and cases not meeting the case definition were excluded. There were issues with case ascertainment for four of the studies examined. Three had concerns with case ascertainment because it was possible that not all isolates were tested for resistance and were, therefore, given a score of ‘1’ because some of the isolates not tested may have met the case definition of resistance (20,26,34). One study had substantial issues with case ascertainment and was deemed poorly protected from bias due to the fact that the study was describing antimicrobial resistance rates, but excluded isolates based on intrinsic resistance and was given a score of ‘0’ (22).

Sampling bias

The weighted kappa statistic was 0.60. Many issues were noted with sampling bias in surveillance systems. Of the 22 studies, the four population-based studies included all patients at risk for antimicrobial resistance; sampling bias was precluded because sampling was not performed and, therefore, were scored ‘2’ for sampling bias (31-34). Seventeen studies scored ‘1’ (13-30). Of these, seven had concerns with sampling because they asked for a certain number of consecutive isolates (13-15,17,19,23,24). Another five studies had sampling issues because they asked for a certain number of isolates with no explicit statement as to the method with which these isolates were to be sampled (18,20-22,28). In addition, five studies scored ‘1’ solely on the basis that the sampling of geographical centres was not reported based on true random sampling (16,26,27,29,30). The study scoring ‘0’ did so because it had no systematic means of selecting participating centres based on the area it sought to measure and samples were obtained from only two laboratories (25).

Multiple counting

The weighted kappa statistic was 0.80. Multiple counting was an area with an important risk of bias. Only eight studies removed duplicate samples and explicitly stated the criteria for doing so and were scored ‘2’ (14-16,19,33,34). There were seven studies that scored ‘1’ because they stated that duplicates were removed, but reported no definition of what constituted a duplicate isolate (17,18,21,23-25,34). Finally, seven studies did not report whether duplicate removal was performed and were given a score of ‘0’ (13,20,22,27,29,31,32).

Laboratory practices and procedures

The weighted kappa statistic for laboratory practices and procedures was 0.60. Laboratory practices and procedures was not a substantial area of concern for bias for most of the studies. Eighteen of the studies had reported thorough protocols, quality-assurance programs, standardized testing and/or centralized testing (13,14,16-24,27-32,34,35). Three of the studies had some concerns with laboratory practices and procedures because documentation was lacking regarding standardized testing among all laboratories or they did not report on any quality-assurance/quality-control programs in place at the laboratories (15,25,33). One study had substantial concerns with a risk for bias because its laboratory methodology was poorly reported (26).

DISCUSSION

In the present study, we identified that contemporary AMR surveillance systems are commonly at risk for bias related to multiple counting and sampling procedures and, to a lesser extent, case ascertainment and laboratory procedures. We did not observe any significant problems with the use of the appropriate denominator data or with case definitions in the studies included.

Multiple counting is a significant potential issue and arises when a case is counted more than once for the same episode of disease (3). While no universal ‘gold standard’ definition exists, it is generally accepted that only the first isolate per patient per episode of disease should be counted (36). Several studies have found that failure to remove duplicates or multiple counting of the same isolates results in an overestimate of both occurrence and rates of resistance (3,37,38). An episode of disease can be based on clinical criteria or on a defined analysis period. In the case of clinical criteria, a second episode is typically defined based on a comprehensive assessment of laboratory and clinical variables such as with repeat illness following complete clinical and/or microbiological resolution of a previous episode. In many cases, particularly with laboratory-based studies, such detailed clinical information is not available and a defined analysis period is used. In these cases, repeat isolates within some time frame (eg, one month, one year) are excluded. Studies have consistently shown that increasing the period of duplicate elimination will reduce the reported incidence and antimicrobial resistance rates (3,37,38).

Sampling bias occurs when the sample under study differs in some systematic way from the larger population of interest (3). One way to minimize or avoid this bias is to include all of the population of interest. However, such population-based studies are often practically difficult to conduct and, in most cases, sampling must be performed (31-34). To be unbiased, a sample should be randomly selected from the overall population of interest. This, however, does not appear to be a common practice in surveillance studies, and convenience sampling from selected laboratories is the usual and potentially highly biased practice. In multicentred studies, hospital-based laboratories – particularly academic tertiary care referral centres – are frequently over represented and, as a result, resistance rates are typically higher than in the population at-large. In addition, the time of day, day of the week, and season of the year may have a significant influence on rates of disease and antimicrobial resistance (38-40). The practice of collecting consecutive samples over a defined period may then be highly influenced by when and where these are obtained.

There are several limitations of the present report that warrant discussion. First, the six biases that we evaluated require, at least to some degree, a component of subjective interpretation, and the possibility exists that other investigators may critique the studies differently. We attempted to minimize subjective interpretation by the use of explicit prespecified criteria for scoring (Appendix). In addition, reviews were conducted independently by two reviewers with generally good or excellent agreement as indicated by the reported kappa scores. Second, our appraisal of study methodology was based on an assessment of methods as reported in the publications. We only reviewed supplemental information surrounding study methodology if it was directly referenced in the index publication under review. Therefore, it is possible that a given study may have been truly protected from a bias, but we assigned a lower score based on a lack of reporting. For example, this is likely the case for the issue of multiple counting with the ABC study (32,41,42). Another possibility is that improvements in methodology not reported in retrieved publications may have been missed by not reviewing all publications from each system. Third, in an attempt to be as systematic as possible, we elected to only evaluate studies on the basis of the six measures of bias that we previously identified (3). There are undoubtedly several other potential biases and considerations that could influence the interpretation of surveillance data that were not included and are not limited to database quality, statistical analysis, and other factors such as timeliness and responsiveness of reporting. Fourth, we only obtained a sample of all current systems by limiting evaluating to all relevant publications in 2008 for practical reasons. In addition, unlike with scoring of studies, the process for selection of systems for inclusion was less systematic and some systems may have been missed. Finally, our overall scores assigned to surveillance systems should not be considered as a linear measure of the quality of study alone because we did not weight the relative importance of the six measures. For example, a study could have a ‘fatal flaw’ in one of the six areas of bias and be considered invalid overall, but potentially still achieve a score of 10/12.
SUMMARY
There are several potential biases that can influence the validity of AMR surveillance information. The potential for bias should be considered in the interpretation and use of AMR surveillance data.

ACKNOWLEDGMENTS: No external funding was received in support of this study. None of the authors has any conflicts of interest to declare.

NOTE: This work was presented in part at the 50th Interscience Conference on Antimicrobial Chemotherapy (ICAAC), September 15, 2010, Boston, Massachusetts, USA.

APPENDIX

SCORING GUIDE
Reviewer: Laupland, Rempel

Reference:

SECTION 1
Inclusion criteria (must have all of the following; check if present)

- Multicentric
- Examine at least one antimicrobial resistant organism
- Laboratory based
- Human
- Ongoing

Context of study:

SECTION 2
Assessment of bias (must select one score for each bias)

1. Denominator data
   0 Denominator data not reported or irrelevant
   1 Used denominator data, but not optimal for stated objective
   2 Optimal denominator data used for objective

2. Case definition
   0 No case definition or inappropriate for objectives/design/reporting
   1 Case definition used, but not clearly appropriate
   2 Case definition matched to objectives/design/reporting

3. Complete ascertainment
   0 Highly likely that cases do not meet case definition may be included; cases that meet the case definition may have been missed or there was no systematic means of case ascertainment
   1 Some cases that do not meet case definition may be included and some cases that meet the case definition may have been missed
   2 All episodes fulfilling case definition included and nonrelevant cases excluded

4. Sampling bias
   0 Arbitrary convenience or non-random sampling; not reported
   1 Sample systematically derived from surveillance population, but at risk for bias in relation to time, space, and/or location
   2 Either population-based or true random sample of surveillance population

5. Multiple counting
   0 Duplicate isolates/episodes not removed or reported
   1 Duplicate isolates/episodes removed, but unclear rationale or explicit criteria
   2 All duplicate isolates/episodes removed with relevant and explicit criteria

6. Laboratory practices
   0 Problems with nonstandardized testing; variable protocols; lacking quality control, testing rules; or not reported
   1 Limitations in consistency, some problems with protocols, quality control
   2 Central laboratory or all laboratories following identical protocol with clear criteria for testing rules, proficiency testing/quality control, and appropriate species level identification

REFERENCES