Evaluation of commercial assays for vancomycin and aminoglycosides in serum: a comparison of accuracy and precision based on external quality assessment

John F. Wilson1*, Alan C. Davis2 and Caroline M. Tobin2

1Department of Pharmacology, Therapeutics & Toxicology, University of Wales College of Medicine, Heath Park, Cardiff CF14 4XN; 2Bristol Centre for Antimicrobial Research and Evaluation, Department of Microbiology, Southmead Hospital, Westbury-on-Trym, Bristol, UK

Received 19 December 2002; returned 18 January 2003; revised 17 February 2003; accepted 23 April 2003

Objective: To compare the accuracy and precision of commercial assay techniques in the measurement of gentamicin, tobramycin, amikacin, netilmicin and vancomycin in serum.

Methods: Data from the measurement of 40 external quality assessment samples from 358 laboratories providing a therapeutic drug monitoring service were analysed.

Results: Significant differences between techniques in accuracy and precision were observed for all drugs. Coefficients of variation ranged from 4.1% to 9.8% for the aminoglycosides and from 6.7% to 11.7% for vancomycin. The percentage difference in measurements from the weighed-in drug concentration ranged from –10.1% to +4.0% for the aminoglycosides and from –3.5% to +5.7% for vancomycin. The Dade Behring Emit immunoassay was notable in producing significantly more outliers (>4 S.D. from the weighed-in concentration) than other techniques in the measurement of gentamicin, amikacin and vancomycin.

Conclusions: All assays performed to a satisfactory standard for measurement in non-renal patients, but none met the more stringent standards desirable for monitoring patients with renal impairment.

Keywords: gentamicin, tobramycin, amikacin, netilmicin, immunoassays

Introduction

The aminoglycosides (amikacin, gentamicin, netilmicin and tobramycin) and the glycopeptide, vancomycin, are first-line antimicrobial agents in the treatment of serious Gram-negative and -positive infections.1,2 In non-renal patients, these antibiotics have short half-lives and were traditionally administered two, three or four times a day. More recently, the recommended aminoglycoside dose has been once daily.3 Although unrelated chemically, all share in common a narrow therapeutic index. Therapeutic drug monitoring (TDM) is recommended in order to avoid ototoxicity and nephrotoxicity, and to optimize antibiotic dosage.4 An assay selected to perform TDM ideally should be rapid, to ensure a fast turnaround time, and be specific, accurate and precise. A number of manufacturers have developed commercial assays in recent years to fulfil these criteria and, despite the relatively high costs of performing immunoassays, they are the current method of choice for most clinical laboratories in the UK.4

For all laboratories providing clinical assays for vancomycin and the aminoglycosides in the UK (and in parts of Europe and the USA), participation in an external quality assessment (EQA) scheme is mandatory. This is not only to fulfil the standards set by the Clinical Pathology Accreditation agency, but also so that individual laboratories can monitor their assay performance against that of their peers. In 2002, there were over 350 laboratories receiving monthly EQA samples from the United Kingdom National External Quality Assessment Scheme (UKNEQAS) for antibiotic assays (Bristol, UK), in collaboration with the Heathcontrol EQA scheme (Cardiff, UK). In this study, we report our findings following a retrospective analysis of EQA data to assess the performance, in terms of accuracy and precision, of assays for vancomycin and the aminoglycosides.

Materials and methods

EQA sample production and distribution

Samples for analysis were prepared from pooled human serum (National Blood Service, Bristol) negative for HIV antibody, hepatitis C antibody and hepatitis B surface antigen. Gentamicin, tobramycin, amikacin (Sigma Chemical Co., Poole, UK), netilmicin (Schering Plough, Welwyn...
Performance of vancomycin and aminoglycoside assays

Table 1. Mean number of measurements per sample by different techniques; in parentheses, the percentage rejected as being >4 S.D. from the weighed-in value

<table>
<thead>
<tr>
<th>Technique</th>
<th>Gentamicin</th>
<th>Tobramycin</th>
<th>Amikacin</th>
<th>Netilmicin</th>
<th>Vancomycin</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abbott TDx</td>
<td>113 (1.1)</td>
<td>48 (1.6)</td>
<td>56 (1.0)</td>
<td>64 (4.6)</td>
<td>105 (0.9)</td>
</tr>
<tr>
<td>Abbott AxSYM</td>
<td>51 (0.9)</td>
<td>15 (1.7)</td>
<td>4 (0.9)</td>
<td>3 (2.8)</td>
<td>48 (0.5)</td>
</tr>
<tr>
<td>Biostat FPIA</td>
<td>52 (1.1)</td>
<td>15 (0.9)</td>
<td>7 (1.8)</td>
<td>41 (0.7)</td>
<td></td>
</tr>
<tr>
<td>Roche FPIA</td>
<td>12 (2.7)</td>
<td>7 (0.8)</td>
<td>10 (3.5)</td>
<td>12 (1.6)</td>
<td></td>
</tr>
<tr>
<td>Sigma FPIA</td>
<td>9 (1.7)</td>
<td>4 (0.7)</td>
<td>9 (1.6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beckman turbidimetric</td>
<td>9 (0.9)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bayer turbidimetric</td>
<td>4 (1.8)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bayer chemiluminescence</td>
<td>3 (0)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dade Behring Emit</td>
<td>8 (4.4)</td>
<td>4 (12.6)</td>
<td>3 (4.9)</td>
<td>6 (6.9)</td>
<td></td>
</tr>
<tr>
<td>CEDIA</td>
<td>4 (1.4)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Abbott TDx and Abbott AxSYM, Abbott Laboratories, Diagnostics Division, Irving, TX, USA; FPIA, fluorescein-polarization immunoassay; Biostat, Biostat Diagnostics, Stockport, UK; Roche, Hoffmann-La Roche, Diagnostics Division, Basel, Switzerland; Sigma, Sigma-Aldrich Co., Poole, UK; Beckman, Beckman Coulter, Fullerton, CA, USA; Bayer, Bayer, Pittsburgh, PA, USA; Dade Behring Eemit, enzyme multiplied immunoassay technique, Dade Behring, San Jose, CA, USA; CEDIA, cloned enzyme donor immunoassay, Microgenics Corp., Fremont, CA, USA.

Garden City, UK) or vancomycin (Eli Lilly, Indianapolis, IN, USA) were added to serum accurately by weight, to give clinically relevant concentrations from subtherapeutic to toxic. Serum was mixed thoroughly and dispensed as 0.5 to 1 mL aliquots. The serum samples were not frozen after antibiotic-spiking, no preservative was added, and they were distributed monthly, by post or courier, within 1 or 2 days of manufacture to EQA scheme members. All antibiotics were distributed as single analyte samples, apart from gentamicin and vancomycin, which were distributed in combination for the last 14 months. Laboratories were requested to report, within a 2–3 week deadline, the antibiotic concentration measured using their routine protocol and the assay technique. Data from the samples distributed in the 40 months between January 1999–April 2002 were included in the statistical analysis.

Data screening

Results for each analyte were first screened to remove transcription and misclassification errors. Measurements >4 S.D. from the weighed-in concentration were rejected using the robust method of Healy. Significantly different (P < 0.05) differences between immunoassays in the frequency of rejected measurements were identified by χ² tests.

Assessment of accuracy and precision

Non-rejected data for each analyte were classified according to technique, and the mean and S.D. calculated for the subgroups. Data from techniques providing a total of <70 measurements for an analyte were omitted from the analysis.

The difference from the weighed-in concentration in the all-technique mean result for each analyte showed an increasing variability with increasing drug concentration. The difference from the weighed-in concentration in the individual-technique subgroup mean, expressed as a percentage, was a more stable parameter with respect to drug concentration and was therefore selected to assess technique accuracy.

The S.D. of measurements increased with increasing drug concentration. Expressing the S.D. as a percentage coefficient of variation (%CV) relative to the all-technique mean (S.D./mean × 100) gave a parameter more uniform with respect to drug concentration, and %CV was therefore chosen to assess technique precision. However, %CVs increase towards infinity as drug concentrations approach zero. For samples containing the lowest drug concentrations, %CVs were large compared with other data, and eight data sets with %CVs from 13–32 were omitted from the analysis. Gentamicin: 0.3 mg/L; tobramycin: 0.2 and 0.3 mg/L; amikacin: 2.0 and 2.5 mg/L; netilmicin: 0.7 mg/L and vancomycin: 3.8 and 4.4 mg/L. In addition, there were comparable large errors in the weighed-in concentration calculated for nine samples, and these data were also omitted from the analysis. The number of samples included in the analyses were gentamicin 38, tobramycin 36, amikacin 38, netilmicin 36 and vancomycin 35. Techniques were compared by one-way analysis of variance, and significant differences between means identified by the Student–Newman–Keuls test (P < 0.05).

Results

The combined membership of the schemes was 358 laboratories; 75% of the laboratories were in the UK, 19% in Europe and 6% outside Europe. The mean number of measurements for each sample by different techniques for the five study analytes are given in Table 1, together with the percentage of measurements rejected as outliers (>4 S.D. from the weighed-in concentration). The Dade Behring Emit technique produced significantly more rejected measurements for gentamicin than the five techniques with ≤1% of outliers. For amikacin and vancomycin, Emit differed significantly from all other techniques. The Roche FPIA produced significantly more outliers than other techniques (except Emit) for amikacin. There were no significant differences between techniques in frequency of outliers for tobramycin and netilmicin (P > 0.5). Inspection of the data for Dade Behring Emit and Roche FPIA, where significantly greater numbers of rejects occurred, showed that samples producing rejected measurements were distributed across the full concentration range studied. There was no clustering of rejected measurements in samples containing low drug concentrations.

The mean ± S.E.M. percentage difference in the technique mean from the weighed-in value, and the mean ± S.E.M. %CV by the different techniques, are displayed in Figures 1 and 2, respectively. Significant differences in both accuracy and precision between techniques were present for all drugs. Subsets of techniques not differing significantly (P > 0.05) are underlined in the Figures. There were no significant interactions (P > 0.05), between technique and drug, in
the measures of accuracy and precision between samples in which gentamicin and vancomycin were alone or combined.

**Discussion**

Significant differences in the precision and accuracy of techniques available for assay of aminoglycosides and vancomycin were demonstrated in measurements in EQA samples undertaken by clinical laboratories that provide a routine TDM service. Spiked EQA samples simulate patient material, but may differ in subtle ways. In this study, the high quality of the matrix (i.e. pooled human serum with no added preservative) provides a close commutability with patient samples. We would expect observed differences between techniques to be replicated with clinical samples. An issue not addressed by the study was technique sensitivity. Data for the lowest drug concentrations were omitted from the statistical analyses, and the comparisons, therefore, refer only to the full therapeutic range.

The popular Abbott assays were in the statistical group with the highest precision for four out of five drugs (exceeded only by the Roche FPIA for tobramycin). The AxSYM assays were slightly superior in precision to those on the TDx, reaching significance for netilmicin. The AxSYM assays were also more precise than Sigma FPIA assays for gentamicin and amikacin, and the Biostat FPIA for amikacin. The CEDIA, turbidimetric and chemiluminescent technologies were of comparable precision to the various FPIA assays for gentamicin. In contrast, the Dade Behring Emit assay differed...
Performance of vancomycin and aminoglycoside assays

There was a difference in accuracy of techniques between the aminoglycosides and vancomycin. For the aminoglycosides, there was a general tendency for techniques to underestimate drug concentration relative to the weighed-in concentration. The latter is a gravimetric estimate from a single laboratory and subject to error. The observed differences in accuracy between techniques are thus to be interpreted in a relative manner. The Dade Behring Emit, Beckman turbidimetric, Biostat FPIA and Sigma FPIA assays produced results closest to the weighed-in value. The Abbott AxSYM, Abbott TDx and Bayer turbidimetric assays produced significantly lower measurements, whereas the Bayer chemiluminescence and CEDIA assays were significantly lower again for gentamicin. The Roche FPIA was unusual in showing biases that differed between

Figure 2. Mean ± S.E.M. percentage coefficient of variation of measurements by different techniques in samples distributed by the UKNEQAS and Heathcontrol EQA scheme for antibiotics between January 1999–April 2002. Lines span techniques that were not significantly different (P > 0.05, Student–Newman–Keuls test). See Table 1 for explanation of technique abbreviations.
analytes. It gave significantly higher measurements for gentamicin and significantly lower values for amikacin than any other technique. The latter negative bias may be responsible for the significantly greater number of rejected measurements identified relative to the weighed-in value for the Roche amikacin FPIA.

For analysis of vancomycin, all methods, with the exception of Abbott AxSYM, gave a positive bias relative to the weighed-in value, with Dade Behring being numerically closest to the weighed-in value. The negative bias of the AxSYM technique differed significantly from all other methods. The explanation for this variation lies in the differences between assays in their cross-reactivity to an inactive crystalline degradation product of vancomycin (CDP-1) produced by spontaneous degradation in samples and found in high concentrations in patients with renal failure.6,7 Assays other than the AxSYM technique are based on polyclonal antibodies that cross-react, to varying degrees, with CDP-1, whereas the AxSYM system uses a murine monoclonal antibody with less cross-reactivity. The AxSYM assay will thus report lower, and potentially more accurate, estimates of vancomycin depending on the CDP-1 content of samples.

Although statistically significant differences were identified between techniques, their scale relative to clinical needs should be considered. Desirable standards of performance can be estimated from the time interval between doses and the average elimination half-life for a drug.8 For a typical aminoglycoside with a 3 h half-life and dosing every 8 h, a %CV of 18 is a realistic analytical goal. This level was achieved by all techniques where the total error resulting from the combination of imprecision and inaccuracy varied between 9%–16%. However, if renal impairment were to increase half-life to 24 h, the desirable %CV falls to 3, and no available technique met this criterion. For vancomycin, a half-life value of 6 h, and dosing every 12 h, equates to a %CV of 15 as a desirable goal. All available techniques met this standard, with an observed total error varying between techniques by 10%–14%. Again, reduced renal clearance will increase the desirable standard of performance beyond that currently achieved. The increase in dose interval that would result from the use of once-daily dosing raises the desirable CV, but not sufficiently to make current assays satisfactory for use in renal patients.

A more worrying aspect of performance was the fact that significant differences between techniques were detected in the number of measurements >4 S.D. from the weighed-in concentration. If these had all been the result of non-assay-dependent errors, such as transcription errors, between-technique differences would not have been expected. The poorer performance of the Dade Behring Emit assays in this comparison correlated with their poorer precision seen in analysis of non-rejected data, and suggests that Emit assays are more prone to producing occasional measurements with sufficient bias to have potential adverse clinical consequences.

In summary, all the commercial immunoassays reviewed by the study achieved the desirable standard of performance required for the provision of a routine therapeutic drug monitoring service for aminoglycosides and vancomycin in normal patients. In patients with renal failure, a higher standard of precision and accuracy is required and no currently available assay met the theoretical standard. Users of the Dade Behring Emit assays should monitor their assays, by means of quality control procedures, for occasional grossly discrepant results, and seek to identify and eliminate the cause.

Acknowledgements
We would like to acknowledge the contribution of Dr Les White to this work. His death in April 2002 prevented his appearance as a co-author of this paper.

References

J. F. Wilson et al.