Analytical goals in therapeutic drug monitoring

LARRY D. BOWERS

The practice of therapeutic drug monitoring is based on several disciplines, including pharmacokinetics, pharmacodynamics, and chemical analysis. The impact of the analysis on the determination of pharmacokinetic parameters is not well appreciated. Analytical goals in therapeutic drug monitoring should be established by determining the nature of the problem to be solved, selecting the appropriate matrix and methodology to solve the problem, and developing valid analytical schemes that are performed competently with appropriate quality and interpreted within the framework of the problem.

Therapeutic drug monitoring (TDM) had its modern foundation in the early 1970s when epileptic patients on phenytoin were demonstrated to have better seizure control and fewer side effects if treated with dosages adjusted according to blood concentrations of the drug rather than with standard doses based on the patient’s weight [1].

This individualization of therapy and improvement in quality of life are at the very heart of the philosophy of healthcare. In the present era of cost-effective health business, considerations of cost have become the primary issue for all aspects of laboratory operation, including TDM. Individualization of therapy with TDM should contribute to cost-effective patient management through detection of noncompliance; detection and elimination of drug side effects; detection of unusual metabolism and adjustment of dosage based on individual metabolism; and detection of unusual metabolism and adjustment of dosage based on the effects of disease. The economic impact of the selection of analytical methodology is widely appreciated. There is also, however, an impact of analytical methodology on the quality and scope of service, the significance of which is not always appreciated.

The development of any analysis is based on the principles summarized in Table 1. For the result of an analysis to contribute to the understanding of the problem, all of these steps must be well thought through. Definition of the problem is a key component in the design of a good analytical scheme. By understanding the way in which the data generated by the analysis are to be used, an analytical approach that is fit for that purpose can be selected. Selecting an appropriate sample or analytical method is an important component of quality testing, but too often either is determined by efficiency rather than considerations of appropriateness or is outside of the laboratory’s control. All too frequently, the focus of the clinical laboratory is on competent performance and quality assurance of the assay and not on the appropriateness of the analysis or its contribution to the patient’s care. Interpretation is the final and key element of any analytical scheme, and it requires knowledge of the strengths, limitations, and information content of all of the components of the scheme. I will discuss the analytical goals of TDM from this perspective.

DEFINING THE PROBLEM
Defining the analytical problem before embarking on development of an assay procedure is a critical step. The fundamental question is the objective of the analysis. The requirements for analytical precision, accuracy, specificity, and limits of detection need to be understood before sample type or analytical method is selected. Frequently, discussion of these issues can illuminate additional problems or concerns that must be addressed before an adequate analytical scheme can be developed. Also part of the consideration is the number of samples, turnaround time, and access to analytical services at specific hours. None of this is new to the clinical laboratorian, and the emphasis on quality specifications in laboratory medicine [2] should be emulated in TDM.

SELECTING THE SAMPLE
Interaction of the laboratory in sample selection and collection is a key component of TDM. This involvement can take several forms. An important problem in therapeutic monitoring of digoxin, especially among outpatients, is taking a blood sample during the first 6 h after dosing. As demonstrated by Bernard et al. [3], working with clinicians to dose patients at night so that blood–tissue equilibrium has been achieved before their morning
Selection of an analytical method requires understanding of the analytical problem. For example, the analytical method selected can have a significant effect on all of the measured pharmacokinetic parameters that form the basis of therapeutic drug monitoring: half-life, volume of distribution, trough concentrations, peak concentrations, and therapeutic interval. The difference in the concentration–time profile resulting from an assay specific for a parent drug and an assay that responds equally to the parent and a long-lived metabolite is shown in Fig. 1. Clearly the slopes of the two lines, which are a measure of the apparent drug elimination half-life, are different. The y-intercept, which is the measure of the volume of distribution, is also different between the two assays. The peak and trough concentrations would also differ, as can be seen from the lack of overlap of the log-concentration vs time curves at any point. This in turn results in differences in the therapeutic intervals that are based on either the peak or the trough concentrations. The differences between the elimination rate of the parent compound and its metabolite(s) have a substantial effect on the magnitude of the effect shown above. Aspesley et al. [6] have recently reported on the effects of methodology on the determination of the pharmacokinetic parameters of a new dosage form of cyclosporine.

Interestingly, the therapeutic intervals cited in textbooks, pocket guides, and laboratory procedures have no indication of the method used to establish the interval, nor cautionary notes regarding the importance of the method. One of the few drugs for which the published therapeutic interval has changed with methodology is quinidine. The original method for quinidine—solvent extraction followed by measurement of fluorescence—detected the parent drug, several metabolites, and the dihydroquinidine that was found in dosage forms; the accepted therapeutic interval was 3–8 μg/L. A more specific double-extraction method with fluorescence detection was developed that detected quinidine, 3-hydroxyquinidine, and quinidine-N-oxide; the therapeutic interval for this assay was 2–5 μg/L. Development of HPLC methodology with fluorescence detection allowed specific detection of quinidine alone, which resulted in a further decrease in the therapeutic interval, to 1.5–4.5 μg/L. The introduction of homogeneous immunoassays for quinidine, with various cross-reactivities with its metabolites, has resulted in the use of the double-extraction/fluorescence therapeutic interval without clear documentation of its validity in this context.

In fairness, however, the guidelines used for determining the limits of therapeutic intervals are not well defined. Review of the literature for theophylline therapeutic interval data indicates that intervals determined with either
the peak or the trough concentrations overlap. Although it is widely assumed that trough concentrations are used to establish the therapeutic interval, it is important for the laboratorian to establish whether peak or trough concentrations were used.

Despite the above emphasis on the effect of the analytical methodology in establishing a therapeutic interval, we should recognize that the cause of this effect is the quantification of different analytes. Pharmacokinetic parameters generally derived in TDM require specific measurement of parent drug or metabolite with no analytical bias. Establishing the condition of nonbias and (or) method equivalence thus becomes an important consideration.

Although textbooks [7–9] and journals [10–12] provide excellent treatments of statistical tools and their application to methods evaluation, the literature remains strewn with poor-quality TDM method comparisons. For example, the effect of the range of concentrations measured on the value for Pearson’s correlation coefficient, r, is well known [10, 11]. Nevertheless, this parameter is frequently used inappropriately to document the agreement between two methods. As Westgard and Hunt stated in 1973, “The correlation coefficient . . . is of no practical use in the statistical analysis of comparison data” [10]. Vine and I have commented [13] on the fact that, although the correlation coefficient for the comparison of a monoclonal immunoassay for cyclosporine with a reference HPLC method over the interval of concentrations from 20 to 3000 µg/L was 0.97, the immunoassay results corresponding to the HPLC concentration at the critical decision point at the upper end of the therapeutic interval could be equal or could differ by as much as a factor of two.

As the minimum effort in a methods comparison, the investigators should evaluate and report the slope and the standard deviation of the slope, the intercept and the standard deviation of the intercept, the standard deviation of the estimate (Sy|x), and the distribution of the residuals as a function of the comparison method value. The slope and the standard deviation of the slope are an indication of a proportional bias and its statistical relevance (ideal slope = 1.0). The intercept and its standard deviation are an indication of any fixed bias and its statistical relevance (ideal intercept = 0.0). Sy|x, a measure of the standard deviation between the measured y values and the y values predicted by the regression line, is independent of analyte concentration range. The residuals should be randomly distributed about 0 and be independent of the concentration of the analyte if a simple linear regression fit is appropriate. If simple linear regression is not appropriate, weighted linear regression methods and methods that allow for error in the comparison method (independent variable, x) should be used [12].

Bland and Altman have suggested plotting the difference between the two methods as a function of the mean value of the two methods to document agreement at the 95% confidence level [14]. Although the Bland and Altman approach has gained wide acceptance, it is of paramount importance to apply the correct statistical approach and test the appropriateness of the approach rather than blindly applying the most frequently used test. In reference to their method, Bland and Altman state, “. . . we fear that instead referees will begin to demand our approach in studies where it, too, is inappropriate” [15].

Another approach to method equivalence is to document the comparability of the effect of the methodology on patient treatment. This approach has been used previously in clinical chemistry to evaluate the quality specification for analytical bias based on diagnostic misclassification [16]. Using the therapeutic interval, preferably determined with the appropriate method if there is bias between methods, one can classify the decisions that should be made on adjustment of therapy based on the analytical results. In an evaluation of cyclosporine method comparability (Bowers, unpublished data), the decision-related data from the same sample assayed by a monoclonal immunoassay or an HPLC method agreed with regard to dosing adjustment in only one-third of the cases. In another third, the HPLC result was in the therapeutic interval whereas the immunoassay indicated a toxic concentration, and thus a decreased dose; and in the final third, the HPLC result suggested a subtherapeutic concentration and the immunoassay gave a result within the therapeutic interval. Clearly, analysis of the effect of a method’s results on patient care has more impact than application of a statistical correlation.

Establishing analytical variation goals for drug assays should include the principle of critical decision limits. Glick [17] suggested that the maximum random error be ±20% of the therapeutic interval at the decision limits for the therapeutic interval. The performance goals adopted by the Clinical Laboratory Improvement Amendments (CLIA 1988) appear to be based on this proposal (Table 2). Burnett and Ayers [18] proposed that analytical error should not prevent differentiation of a drug concentration within the therapeutic interval whereas the immunoassay indicated a toxic concentration, and thus a decreased dose; and in the final third, the HPLC result suggested a subtherapeutic concentration and the immunoassay gave a result within the therapeutic interval. Clearly, analysis of the effect of a method’s results on patient care has more impact than application of a statistical correlation.
midway between 0 and the lower decision limit of the therapeutic interval from either of those concentrations. Their concept resulted in an analytical variation that was one-sixth of the concentration at the lower end of the therapeutic interval.

Fraser [19, 20] proposed analytical variance goals based on the intraindividual variability model of Harris [21], but selected the interval from peak to trough as the biological variability. The analytical variability ($SD_a$) is thus given by

$$SD_a < 1/4(2^{T/\tau} - 1)/(2^{T/\tau} + 1) \times 100\%$$

where $T$ is the dosing interval and $\tau$ is the elimination rate constant. Thus drugs with small dosing intervals or long half-lives required better precision. This approach is flawed in that the analytical goal may be to adjust the dosage for an individual and by the assumption of intravenous administration of the drug. The acceptable analytical variability should be based on its impact on the measured quantity, and therefore analytical goals for detecting noncompliance should be different from those for detecting interindividual differences in elimination half-life. If one accepts that an analysis begins with collection of the samples, then even accurate collection times must become part of the analytical goal in the case of half-life determinations. None of the models to date consider the impact of preanalytical variables on outcome.

Jenny [22], in evaluating analytical goals for therapeutic monitoring of theophylline, proposed two methods “fit for purpose” for two of the goals of TDM: decision limits and dosage adjustment. He proposed that the diagnostic efficiency at a decision value, such as the upper limit of the therapeutic interval, should not be degraded by >5%. From the estimates of biological variability of theophylline, this required an imprecision (CV) at the decision limit of 2.8%. Harris and Boyd [23], reevaluating his work, found that an analytical CV of 12.5% should be acceptable. Jenny also recognized that dosage adjustment is a goal of TDM that may require different analytical precision. Using a simple one-compartment model, he proposed that the analytical CV be such that the predicted dosage would fall within the therapeutic range with 95% confidence. In Jenny’s calculations, this required an analytical variability between 2.5% and 18%, depending on the clearance. Harris and Boyd’s [23] extension of the work of Jenny suggested that a CV of 7% is adequate.

The development of immunosuppressive drugs has presented a unique opportunity for laboratorians. Rather than attempting to set analytical goals after a drug is in clinical use, investigators have developed recommendations that have been established and adopted at several consensus conferences on therapeutic monitoring for immunosuppressive drugs [5, 24, 25]. Despite these recommendations, however, analytical systems that do not meet the criteria are allowed to enter the field and are in use, primarily because of cost constraints in the healthcare system. We do not know the incidence of morbidity related to inappropriate dosage adjustments resulting from use of different cyclosporine methods that do not give equivalent results, but the number of affected patients must be in the thousands. Additional opportunities to set analytical goals and develop assays for recombinant proteins and anti-sense DNA pharmaceuticals should arise.

**METABOLITE REPORTING**

One of the most important and difficult problems facing TDM is the issue of metabolite reporting. For example, my colleagues and I reported on an incidence of quinidine toxicity that was correlated to an increased 3(S)-3-hydroxyquinidine (3-OHQ) concentration [26]. The patient exhibited low concentrations of quinidine as determined by a specific HPLC method, which the clinicians attributed to poor absorption. Because HPLC was being used, we were able to observe an unusual profile of metabolite peaks (Fig. 2) in which the 3-OHQ concentration was several times greater than that of the parent compound. We related this observation to the clinicians, but could provide no clear interpretation of it. In a prospective study then undertaken, 3% of patients on quinidine had similar results with respect to metabolism and toxicity, and an additional 2% of patients had increased 3-OHQ/quinidine ratios but no apparent toxicity.

This case illustrates several of the difficult issues in reporting metabolite concentrations. The literature relating metabolite concentrations to toxicity, particularly when there are multiple metabolites as well as the parent compound, is almost nonexistent. Second, in most circum-

---

**Fig. 2.** HPLC chromatogram of a “normal” metabolic pattern for quinidine (A) and the pattern observed for the index case (B).

Note the differences in the concentrations of quinidine-10,11-diol (1 and 2), and 3-hydroxyquinidine (3) at the same relative concentration of quinidine-N oxide (4) and quinidine (5). Peak 6 is the internal standard, quinine. Other reported metabolites of quinidine were not detected. The quinidine concentration (peak 5) in the samples was 2.1 mg/L (A) and 1.5 mg/L (B). Reprinted from Bowers et al. [26] with permission.
stances in which metabolite activity has been studied, the drug has been studied alone, without the mixture of compounds that might be observed in a patient's sample. Thus the clinical chemist has the continued dilemma of how to report and interpret the presence of metabolite concentrations or the total drug activity. Despite the observation of 3-OHQ-related toxicity, immunoassay procedures have been introduced as recently as 1995 that have considerable cross-reactivity with metabolites. The issue of inaccurate, and potentially inconsistent, assessment of parent vs metabolite concentrations in these kits was not addressed.

STANDARDIZING THE ASSAY AND INTERPRETING THE RESULTS
Method validation is becoming a more universally important consideration, both for TDM and for toxicology. The pharmaceutical industry has mounted a worldwide effort to harmonize the concepts used in validation, which are summarized in Table 3 [27]. The concept of reproducibility in this system refers to the ability to achieve the same results between laboratories. The issue of specificity should clearly be of concern to clinical laboratories. The ability to differentiate metabolite and parent compound is important in many cases of pharmacokinetic as well as TDM applications.

It is important not to confuse specificity with analytical accuracy. In the discussion of analytical goals above, accuracy was assumed and only precision goals were considered. Recent experiments in our laboratory using electrospray HPLC/MS/MS methodology for steroid glucuronides (Borts and Bowers, unpublished data) produced ion suppression that varied by a factor of 8–10 between samples and also varied between various analytes at different retention times. Two lessons emerge here: High technology does not guarantee accuracy, and sample preparation can have an important effect on even the most sophisticated analytical tool. Fortunately, we were able to use deuterated internal standards for each of the compounds of interest to minimize the effect of the signal suppression.

The concept of assay robustness is relatively new to the clinical laboratory. To establish a robust assay, the effect of each component of the analysis on the results must be evaluated. In the case of chromatographic methods, this would include the effect of column lot, the effect of small changes in eluent composition, and the effect of temperature on the reproducibility of retention times and quantitative data. This should become the standard for anyone publishing analytical methods. Finally, the determination of limits of detection is an important part of assay validation. Although many analytical schemes determine the limit of detection from the signal-to-noise ratio of a blank, the role of the analytical matrix in determining the limit of detection must also be evaluated.

SUMMARY AND CONCLUSIONS
One of the key issues for TDM is when in the course of drug development the basis for therapeutic monitoring is established. Clinical trials based on dosage by body weight delay the evaluation of the beneficial aspects of TDM. In recent years, the Food and Drug Administration appears to have recognized this fact and has been more proactive in assuring that accurate, relevant monitoring is performed and that dosage adjustment based on monitoring is included in phase II clinical trials. The development of reference analytical methods that can achieve the necessary accuracy and precision at critical decision points must precede initiation of monitoring. Although this approach may have a large economic impact on the pharmaceutical company involved, it is the only method that ensures an accurate determination of TDM goals.

Among the problems that remain to be resolved is our failure to deal with the reporting of metabolite and parent concentrations in a clinically useful form. The inability to clearly interpret and communicate the impact of metabolites weakens the value of monitoring to the physician. Similar difficulties can be documented with drug–drug interactions. Studies documenting the economic benefit of TDM are scarce, but studies of the impact of analytical goals are nearly nonexistent. Unfortunately, few funds have been available for this type of work, and the future does not bode well for increased support.

Finally, TDM must reexamine the relative importance of pharmacokinetic and pharmacodynamic parameters in monitoring. Many drugs have effects that are but poorly related to the physiological fluid concentration, particularly when there is a delay between drug administration and clinical effect. Examples of this include viral load monitoring for AIDS treatment, tumor load assessment for chemotherapeutic agents, and bone marker changes for drugs used to treat osteoporosis. New analytical challenges, based on definition of the problem, will continue to arise. TDM and its analytical component should remain a dynamic part of healthcare.

---

**Table 3. Method validation issues.**

<table>
<thead>
<tr>
<th>Accuracy</th>
<th>Precision</th>
</tr>
</thead>
<tbody>
<tr>
<td>Specificity</td>
<td>Limit of detection</td>
</tr>
<tr>
<td></td>
<td>Limit of quantitation/identification</td>
</tr>
<tr>
<td></td>
<td>Linear dynamic range</td>
</tr>
<tr>
<td>Reproducibility</td>
<td>Repeatability</td>
</tr>
<tr>
<td>Robustness</td>
<td></td>
</tr>
</tbody>
</table>

---

**References**