Understanding and Reducing the Experimental Variability of *In Vitro* Plasma Protein Binding Measurements

HAIPING WANG,[1] MATT ZRADA,[2] KEN ANDERSON,[2] RAVI KATWARU,[1] PAUL HARRADINE,[1] BERNARD CHOI,[1] VINCE TONG,[3] NATASA PAJKOVIC,[3] RALPH MAZENKO,[2] KATHY COX,[1] LUCINDA H. COHEN[1]

[1]Department of Pharmacodynamics, Pharmacokinetics and Drug Metabolism, Merck Research Laboratories, Rahway, New Jersey
[2]Department of Pharmacodynamics, Pharmacokinetics and Drug Metabolism, Merck Research Laboratories, West Point, Pennsylvania
[3]Department of Pharmacodynamics, Pharmacokinetics and Drug Metabolism, Merck Research Laboratories, Boston, Massachusetts

Received 16 May 2014; revised 9 July 2014; accepted 11 July 2014

Published online 12 August 2014 in Wiley Online Library (wileyonlinelibrary.com). DOI 10.1002/jps.24119

ABSTRACT: The experimental measurement of plasma protein binding is a useful *in vitro* Absorption Distribution Metabolism and Excretion(ADME) assay currently conducted in both screening and definitive early development candidate modes. The fraction unbound is utilized to calculate important pharmacokinetic (PK) parameters such as unbound clearance and unbound volume of distribution in animals that can be used to make human PK and dose predictions and estimate clinically relevant drug–drug interaction potential. Although these types of assays have been executed for decades, a rigorous statistical analysis of sources of variability has not been conducted because of the tedious nature of the manual experiment. Automated conduct of the incubations using a 96-well equilibrium dialysis device as well as high-throughput liquid chromatography–mass spectrometry quantitation has now made this level of rigor accessible and useful. Sources of variability were assessed including well position, day-to-day, and site-to-site reproducibility. Optimal pH conditions were determined using a design of experiments method interrogating buffer strength, CO_2% and device preparation conditions. Variability was minimized by implementing an in-well control that is concurrently analyzed with new chemical entity analytes. Data acceptance criteria have been set for both the in-well control and the range of analyte variability, with a sliding scale tied to analyte-binding characteristics. © 2014 Wiley Periodicals, Inc. and the American Pharmacists Association J Pharm Sci 103:3302–3309, 2014
Keywords: robotics; mass spectrometry; HPLC; protein binding; preclinical pharmacokinetics; equilibrium dialysis

INTRODUCTION

Beginning in 2007, standardized workflows and procedures have been adopted across all Merck bioanalytical laboratories; site-specific protocols have been replaced with a single analytical procedure for each assay. This new model has increased efficiency, and more importantly, flexibility by allowing work to be shared across all of the laboratories in the network regardless of the origination of the request. Standardized workflows provide the framework for an efficient and lean operational model, where work can be easily shifted from laboratory to laboratory. Highly detailed protocols with clearly defined assay parameters are needed. In a discovery environment, a fit-for-purpose strategy often determines experimental design and analytical acceptance criteria.[1] Many factors are taken into account, including sample throughput, resource allocation, and most importantly, how the data are used. Acceptance criteria provide a measure of acceptable analytical error, generally determined by the level of precision and accuracy of calibration standards and/or quality control samples.[2] It is important, however, to understand that certain assays are more sensitive to environmental conditions and, in those cases, interlaboratory variability needs to be carefully assessed, minimized, and controlled.[3]

The determination of the unbound concentration of drugs in plasma, plasma protein binding (PPB), is an experiment that is highly sensitive to assay conditions and presents a serious challenge to bioanalytical laboratories.[4] The measurement of PPB is an important *in vitro* Absorption Distribution Metabolism and Excretion (ADME) assay currently conducted at Merck in both screening and definitive early development candidate selection mode. The objective of the assay is to determine the extent of binding a drug candidate exhibits to plasma constituents, primarily plasma proteins.. Specifically, the unbound fraction of a compound ($f_{u,\ plasma}$) is calculated by taking the ratio of the measured unbound drug concentration over the measured total drug concentration, which may also be reported as a percentage. According to the unbound drug hypothesis, pharmacological activity is determined by the concentration of unbound drug at the site of the therapeutic target.[5] As a result, pharmacokinetic (PK) analyses, PK/pharmacodynamic (PK/PD) models, human PK, and drug–drug interaction predictions generally rely on unbound, rather than total drug concentration in their calculations.[6,7] As a compound may be differentiated and selected for further development based on parameters calculated using PPB data, it is therefore critical that the assay provides the appropriate amount of rigor (accuracy, precision, and reproducibility) so that valid data-driven decisions are made. Experimental variability (intraday and interday) should be determined and controlled, especially in the measurement of highly bound drugs where uncertainty in fraction unbound is highly sensitive to experimental error.[8]

Equilibrium dialysis is a traditional method with a history of widespread use and is our current method for measuring PPB in discovery and early development programs. The use of a commercially available device in a 96-well format facilitates

Correspondence to: Lucinda H. Cohen (Telephone: +732-594-3745; Fax: +732-594-0000; E-mail: lucinda_cohen@merck.com)

Journal of Pharmaceutical Sciences, Vol. 103, 3302–3309 (2014)

robotic automation and increases sample throughput.[9] Because of the widespread use of equilibrium dialysis, common pitfalls and sources of assay variability are the subject of many papers in the literature. Insufficient equilibration,[9,10] volume shift,[11,12] lack of pH control,[13–16] and protein leakage[17] are known to contribute to assay variability and should be controlled during a PPB experiment.

In this study, we report how Six Sigma Methodology[18] was used to identify, reduce, and control variability of the PPB assay. Measurement systems analysis (MSA)[19] and design of experiment (DOE)[20] methods provided a statically rigorous framework for the evaluation of assay variability. MSA interrogates the interday and laboratory ruggedness and repdocibility of the assay, whereas DOE allows multiple factors, that is, sources of variability to be simultaneously and efficiently evaluated. Results of the MSA revealed that small differences in the procedures of different laboratories were contributing to interlaboratory variability; in some cases, experimental protocols were written with an acceptable range for a specific assay parameter instead of an absolute value. Results of the MSA, more importantly, identified the major source of variability of the assay. Lack of pH control was also identified as a source of variability and was investigated by the simultaneous optimization of three incubation parameters by DOE methodology. We established systematic acceptance criteria and the use of in-well controls to monitor assay performance to both increase the data quality and reduce the need to repeat experimental determinations.

EXPERIMENTAL

Chemicals and Reagents

Test compounds warfarin, clozapine, diltiazem, diclofenac, fluconazole, fluoxetine, and verapamil were obtained from Sigma–Aldrich (Madison, Wisconsin). HPLC grade water and all analytical organic solvents used were purchased from Fisher (Fair Lawn, New Jersey). Control sodium ethylenediaminetetraacetic acid human plasma was purchased from BioReclamation Inc. (Hicksville, New York). Chemicals (Na_2HPO_4, NaCl, and $NaH_2PO_4{\cdot}2H_2O$) used in buffer preparation were purchased from Thermo Fisher (Fairlawn, New Jersey). Carbon dioxide gas used during incubations was from Airgas (Meadville, Pennsylvania).

Buffer Preparation

Sodium phosphate (100 mM) and 150 mM NaCl buffer (phosphate buffer solution, PBS) were prepared by following method. A basic solution was made by dissolving 14.2 g/L Na_2HPO_4 and 8.77 g/L NaCl in deionized water. An acidic solution was made by dissolving 15.6 g/L $NaH_2PO_4{\cdot}2H_2O$ and 8.77 g/L NaCl in deionized water. The basic solution was then titrated with the acidic solution to pH 7.40.

Instrumentation

A Hamilton Star Plus liquid handling robot was used to prepare the plates for incubation, aliquotting of buffer and plasma samples after incubation, matrix matching of buffer and plasma, and standard curve preparation. Mass spectrometric analysis was performed on an Applied Biosystems Sciex API 4000 triple quadrupole mass spectrometer (Toronto, Canada) equipped with an ESI source. Analyst software (V. 1.5) was used for data acquisition and peak integration. Ultra-performance liquid chromatography (UPLC) was performed on a Thermo Fischer Scientific (Chelmsford, MA) Transcend dual inlet upltra pressure liquid chromatograph. Waters Acquity UPLC™ HSS T3 (2.1 × 50 mm^2, 1.8 μm) columns were obtained from Waters Corporation (Milford, Massachusetts). Deep 96-well collection plates were purchased from Analytical Sales and Services (Pompton Plains, New Jersey). Plasma pH was measured with a microcombination pH microelectrode (Microelectrode Inc., Bedford, New Hampshire). A Thermo Scientific CO_2 incubator with temperature control was utilized during equilibrium dialysis Statistical analysis was conducted using Minitab version 16.2.1. MSA results were evaluated using gage reproducibility and repeatability, analysis of variance (ANOVA), nonparametric Kruskal–Wallis, and Moods' median tests as well as general statistical tests. DOE methods utilized a two-level full factorial design with three factors, eight runs, and singlet replicates representing the mean of 10 measurements for each run. Main effects including estimated coefficients and two- and three-way interactions were analyzed. All derived terms were free of aliasing.

Dialysis Method

Plasma-spiking solution was prepared by transferring 5 μL of 2 mM dimethyl sulfoxide solution to 995 μL of the plasma to create an intermediate concentration of 10 μM. The intermediate solution was further diluted by taking 100 μL of the 10 μM solution and adding it to 300 μL of plasma yielding a 2.5-μM analyte concentration in plasma. Dialysis was performed using a reusable 96-well HT dialysis micro equilibrium device by HT Dialysis LLC (Groton, Connecticut) and dialysis membrane strips (molecular weight cutoff 12 – 14 kDa). Membrane strips were hydrated by soaking in ethanolic PBS. The dialysis block was prepared by adding 120 μL of the pH 7.4 100 mM PBS to one side of the membrane. Next, 120 μL of the analyte-spiked plasma was added to the other side of the dialysis membrane. After aliquotting into the Teflon block, samples were incubated at 37°C in a humidified incubator for 4 h with 5% CO_2. Following a 4-h incubation period, a 50-μL aliquot was removed from the plasma side of the equilibrium dialysis block and added to a 96-well plate. To this sample was added 50 μL PBS and 200 μL of solution containing 200 nM diclofenac, 200 nM labetalol, and 100 nM imipramine in acetonitrile (acetonitrile internal standard mix). Additionally, a 50-μL aliquot was removed from each well of the buffer side of the equilibrium dialysis block, and added to a 96-well plate. To this sample was added 50 μL plasma and 200 μL acetonitrile internal standard mix solution.

Ultra-Performance Liquid Chromatography–Tandem Mass Spectrometry

Ultra-performance liquid chromatography was performed on a Waters Acquity UPLC™ HSS T3 (2.1 × 50 mm^2, 1.8 μm) column. The samples were eluted from the column at 0.75 mL/min with a stepwise procedure. Mobile phase A was water with 0.1% formic acid and mobile phase B was acetonitrile with 0.1% formic acid. A gradient elution program was utilized where the solvent composition was held at 5% B for 0.25 min and then changed from 5% B to 95% B in 1.5 min. The mobile phase composition was then held at 95% B for an additional 0.4 min. The column was re-equilibrated at the original solvent composition

for 1 min. The total run time was 3.5 min. An Applied Biosystems Sciex API 4000 triple quadrupole tandem mass spectrometer equipped with an electrospray source was operated in the positive ion mode using multiple reactions monitoring. The ion spray voltage was set to 4.5 kV and the auxiliary gas temperature was maintained at 500°C. Nitrogen was used for GAS 1, GAS 2, curtain, and collision gas. The mass resolution was set to a peak width of 0.7 mass units at half-height for both Q1 and Q3. The electron multiplier was set at 2000 V.

RESULTS AND DISCUSSION

Initial Control Assessment

To understand the inherent variability in the experiments being conducted at different laboratories, MSA methodology was employed,[19] whereby the existing variability of the scientists, original assay conditions, and instruments utilized in the experiment could be determined. MSA allows the establishment and differentiation of variability arising from the measurement system (liquid handling robotics, mass spectrometers, and equilibrium dialysis devices) rather than the overall process variability. Warfarin and clozapine were chosen to be test compounds in human plasma representing low unbound fraction (warfarin) and moderate unbound fraction, pH sensitive (clozapine) analytes.[16,21] The same scientist conducted the experiment on 3 separate days using the same instrument at each site for two Merck laboratories. Because of serious concerns about the effect of well location within the HT Dialysis device because of potential location-based leaks or tears, variability was assessed across the entire 96-well device for each analyte. In total, 576 measurements were obtained for each compound. This experiment is conducted early in the drug discovery process as a screening exercise. To that end, the incubation conditions were set at 4 h. Although some compounds may require additional time to equilibrate, if a compound progressed into early development, this screening PPB experiment will be repeated with more rigorous experimental design and conduct, including assessment of the time required to reach equilibrium. The 4-h incubation time also offers the advantage of preventing volume shift because of water evaporation from the dialysis device, as recommended by the manufacturer. The equilibration time of warfarin and clozapine was assessed by conducting incubations at 2, 4, and 6 h. No significant difference in fraction unbound was observed at 6 versus 4 h. It should be noted that this may not be true for other analytes and may be a source of variability.

The results of the MSA experiments are summarized in Figure 1. For the clozapine data shown in Figure 1a, the median unbound fraction obtained was 0.1274 with 95% confidence interval (CI) of the median ranging from .0.1137 to 0.1184. The median warfarin unbound fraction was 0.01307 with 95% CI of the median ranging from 0.01062 to 0.01133. The warfarin results were comparable to previous literature reports for human PPB, whereas the clozapine was somewhat higher than the reported $f_u = 0.05$ value.[22] The two Merck sites utilized different incubation conditions that would ultimately generate different pH environments, under which clozapine has been shown to generate variable f_u results.[14]

Figure 2 shows the boxplot distribution over 3 separate days and two sites for clozapine (Fig. 2a) and warfarin (Fig. 2b). Ninety-six measurements were obtained on each day, using different devices each day. The clozapine dataset represented reasonable reproducibility and repeatability across days and sites, although the occasional spurious value at 0.4–0.6 unbound fraction was observed. In contrast, significantly more outliers were observed across both sites and days for warfarin. In addition, outliers observed for warfarin significantly impacted the mean value determined per day, which was deemed unacceptable.

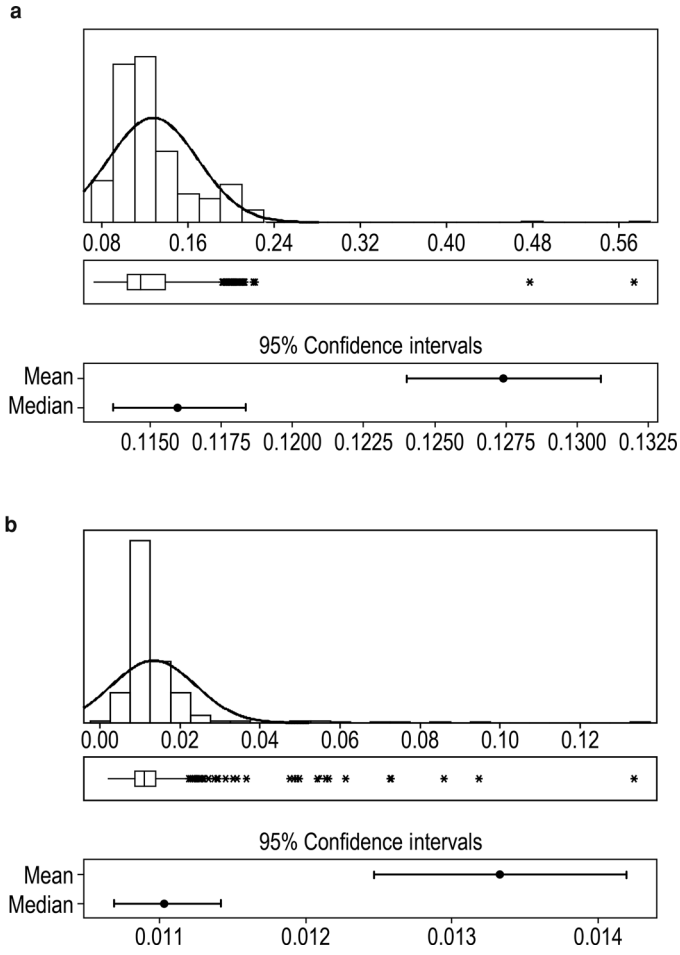

Figure 1. Graphical summaries of human PPB clozapine (a) and warfarin (b) data. Each graph represents 288 individual unbound fraction measurements.

Once the baseline variability was established, the next step consisted of analyzing the warfarin and clozapine data in great detail to determine the sources of variability. In addition to well location, several other parameters were investigated including analyte peak area, internal standard peak area, and plasma and buffer drug concentration determinations. For the 1152 measurements for clozapine and warfarin, 60 outlier values were observed (~5% error rate). Results of this analysis are shown in Table 1. Almost half of the outliers were derived from unusually high drug concentrations in the buffer. Unusually high drug concentration was defined as values more than three standard deviations greater than the overall mean value. Other lower frequency errors included unusually high or low drug concentrations in plasma and no observed plasma or buffer concentrations. In addition, four errors were categorized as no internal standard response, likely because of an error either during sample preparation or LC–MS analysis. Ultimately, the

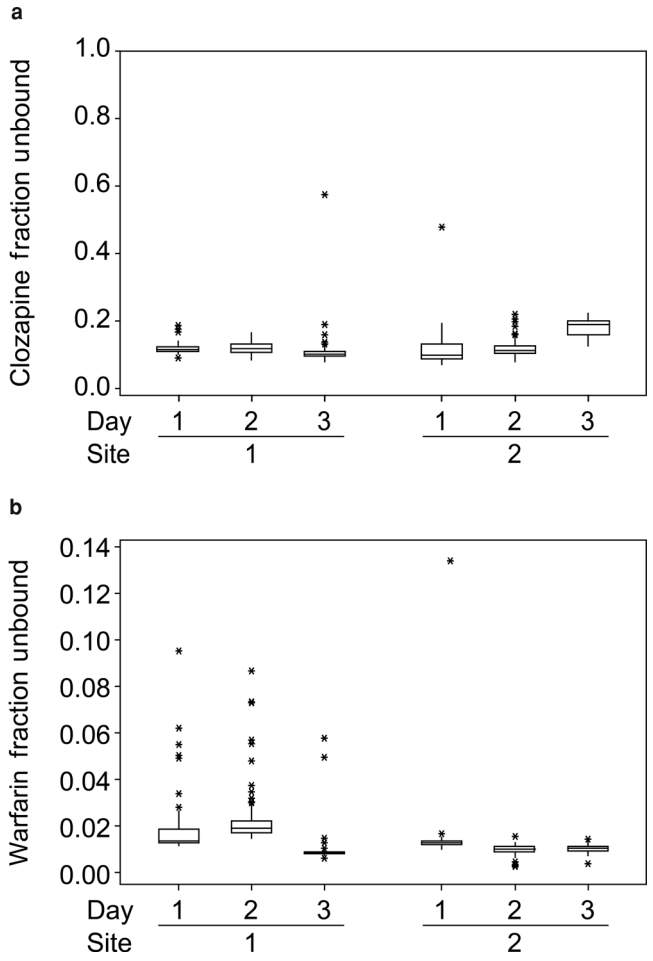

Figure 2. Measurement systems analysis reproducibility and repeatability comparison across sites and days for (a) clozapine and (b) warfarin human PPB.

Table 1. Measurement Systems Analysis Error Types

Error Type	Total Errors (%)	Number of Observations
High buffer concentration	56.6	34
High plasma concentration	5.0	3
Low plasma concentration	26.6	16
No plasma or buffer concentrations	5.0	3
No internal standard	6.7	4
Total		60 errors/1152 incubations

outlier drug concentrations in buffer data were attributed to physical damage to the dialysis membrane from leaks or physical deformation over the course of the experiment. This was determined by photographing the equilibrium dialysis apparatus after incubation and tracing back the unusually high buffer concentration measurements to the well location. Inspection of the apparatus photographs revealed a torn, bent, or crimped membrane in that specific well. This physical damage may be caused by contact from the pipette tips with the membrane. Of the 60 outliers examined, no correlation was observed for any of the outliers with respect to well position in the 96-well plate.

Table 2. pH Design of Experiments ($n = 12$)

Run Order	Cover	CO_2 (%)	Buffer Concentration	Buffer pH	Plasma pH
1	Y	10	100	7.5	7.5
2	N	10	100	7.2	7.2
3	Y	10	10	7.9	7.8
4	N	10	10	7.4	7.2
5	N	5	10	7.8	7.7
6	N	5	100	7.6	7.6
7	Y	5	100	7.6	7.6
8	Y	5	10	7.9	7.8

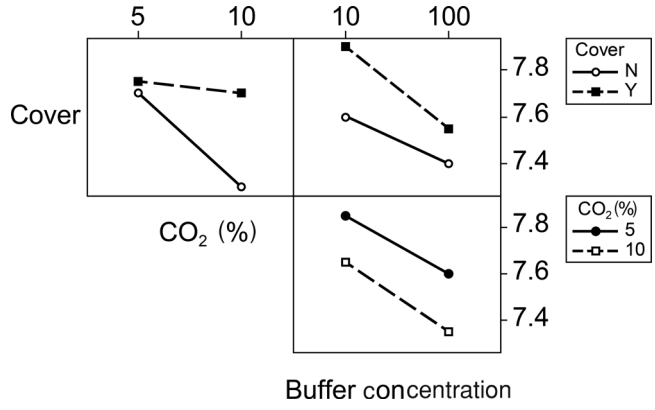

Figure 3. Variable interaction plots for buffer pH.

pH Design of Experiments

In addition to general concerns surrounding the instrumental capabilities of the automation and mass spectrometer, another hypothesized contributor to variability was improper pH control during the experiment. The most prevalent variables that could potentially impact pH were discussed and prioritized to focus on the percentage of carbon dioxide utilized during incubation, the buffer strength utilized during dialysis, and whether or not the equilibrium dialysis device was covered or uncovered during incubation. These three variables to optimize pH were simultaneously interrogated using design of Experiments (DOE) methodology.[20–23]

A simple screening DOE paradigm was utilized, with low and high values for each variable consisting of 10 or 100 mM PBS, 5% or 10% CO_2 and covered or uncovered devices. In total, pH determinations were conducted for eight different combinations of the three variables. Twelve replicate measurements of pH were obtained for both human plasma and buffer for each combination of variables after 4 h incubation. The DOE conditions and resultant pH are shown in Table 2. In general, the standard deviation of the pH measurements was 0.1 pH unit. The most significant factor that impacted the experimental pH was device covering. Figure 3 shows the summary interaction plot for buffer strength, CO_2 percentage, and covering for buffer samples. Similar results were obtained for plasma pH determinations (results not shown). Within each quadrant, parallel lines such as those observed for buffer concentration verses CO_2 percentage indicate no major interaction that impacts pH. In contrast, a significant interaction was observed between CO_2

percentage and covering. This was not surprising, considering that covering the device will prevent the CO_2 from lowering the pH of the plasma and buffer within the device. The identification of the interaction between CO_2 percentage and covering the device also verified the utility of the DOE analysis. Ultimately, the conditions that yielded pH closest to physiological conditions were determined to be 5% CO_2 with no device covering and 100 mM buffer concentration. Under these conditions, the pH was best maintained for the time period necessary for the automation to pipette samples out of the equilibrium dialysis device. Although the pH of the plasma and buffer samples was slightly higher than physiological (7.6 vs. the desired 7.4), maintaining the pH under ambient conditions was significantly easier than with the conditions of 10% CO_2, 10 mM buffer, and no covering, which drifted to pH 7.0 in the 30-min necessary for sample pipetting. In addition, these conditions are consistent with those previously reported in the literature,[14] albeit for 6 h incubation times.

Variability Assessment and Error Proofing

Once the major sources of error from the MSA data and the experimental conditions to best control pH had been identified, the experimental protocol was revised to generate higher quality data. Considering that the most significant source of error was anomalously high drug concentrations in buffer resulting from sample liquid transfer problems or membrane failure, the relatively low frequency of this error is best monitored by the use of an experimental control in each well of the device, combined with the analyte for each measurement. Warfarin was chosen as the in-well control because of its extensive characterization. Although warfarin binds extensively to plasma proteins, its relatively low concentration relative to the approximately 500 mM HSA concentration obviated any concerns about competitive binding with the analyte of interest. The significant impact of outliers on the observed value for warfarin also supports its choice as an in-well control. In addition to instituting the warfarin in-well control, the number of replicates for each analyte was increased from three to six to provide the ability to accept or reject individual analyte values and still report data.

Considerable debate has occurred within the drug metabolism and pharmacokinetics community regarding the definition of acceptable variability in PPB measurements. Although the stage of discovery or development obviously impacts this, generally the most popular and mathematically rigorous proposed range is twofold values, although this is not universally deemed appropriate. However, twofold variability for highly bound values (<0.001 unbound fraction) is generally viewed as unacceptable. For example, if a compound was observed to have values ranging from 0 (not detectable) to 0.002 fraction unbound, this amount of variability would be deemed unacceptable and the experiment repeated. Another confounding consideration is the context in which PPB is leveraged, whether for calculation or prediction. Critical decisions such as calculation of human safety margins from preclinical data clearly require higher confidence in the absolute data values as well as accuracy and precision.

An internal assessment of model compounds in the low, moderate, and highly bound compounds (results not shown) led to the proposal to assess individual PPB values against the range of values obtained for six replicates. Figure 4 shows a plot of the acceptance limits set by observed mean unbound fraction.

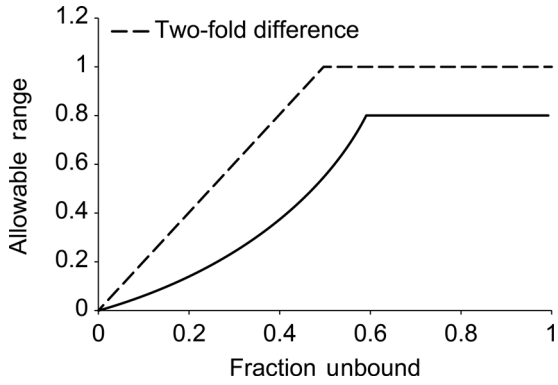

Figure 4. Acceptance criteria: allowed variation as a function of unbound fraction in plasma.

Table 3. Protein-Binding Acceptance Criteria

Mean Unbound Fraction	Mean Unbound (%)	Acceptable Range of Replicates
<0.001	0–0.5	0.0030
0.05	5	0.0320
0.15	15	0.1003
0.35	35	0.3000
0.60	60	0.8000

For compounds with unbound fraction of 0.6–1, the acceptable range for the maximum to minimum values is 0.8. For example, for a dataset consisting of 0, 0.2, 0.4, 0.45, 0.50, and 0.80 unbound fraction, all values would be deemed acceptable and a mean value of 0.39 would be reported. As the unbound fraction decreases, the acceptance limits decrease in a curvilinear fashion. This curve is defined by a set criteria at the middle and upper points for binding. At very low free fraction values (<0.001), the acceptable range is 0.003. At 0.35 unbound fraction, the acceptable range of values is 0.30. A general hyperbolic line described by $y = B+(A/x)$ through these points above gives values of $A = 55.16$ and $B = -54.86$. Thus, more rigorous acceptance criteria have been established for lower free fraction compounds. Table 3 shows representative acceptance criteria.

For each experiment conducted, both the analyte and the warfarin data are evaluated. A minimum of four reportable values are required for results reporting for a valid result. If fewer than four values are satisfactory, then the experiment for the compound fails. If individual wells fail for the warfarin control, then those wells also fail for the compound. Additional data quality assessments are conducted and include assessment of matrix ion suppression, chromatographic retention, and mass spectrometric response.

The dotted line in Figure 4 shows the theoretical acceptance criteria range using the "twofold rule." The calculated maximum acceptable range is reached at an unbound fraction of 0.5, as the hypothetical free fraction cannot exceed 1.00. At across the entire range of free fraction values, the Merck acceptance criteria offer a more stringent threshold than the twofold rule of thumb.

Control Data and Assay Performance Assessment

After standardization of assay conditions and institution of the previously described acceptance criteria, the performance of the assay was monitored for several months. Figure 5 shows the

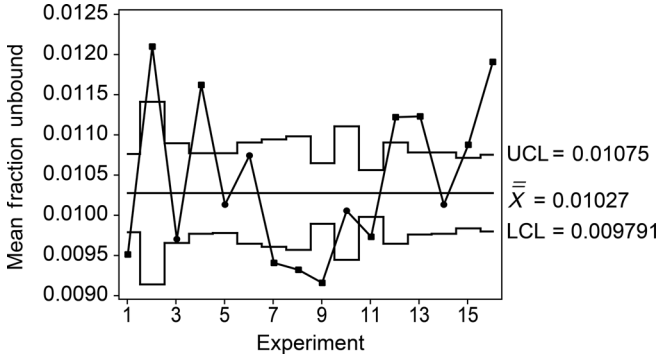

Figure 5. Day-to-day variability of warfarin in-well control for rat plasma determinations. Individual datapoints represent daily mean value. The center line represents the grand mean. Daily confidence limits are plotted against the grand mean. UCL, upper confidence limit; LCL, lower confidence limit.

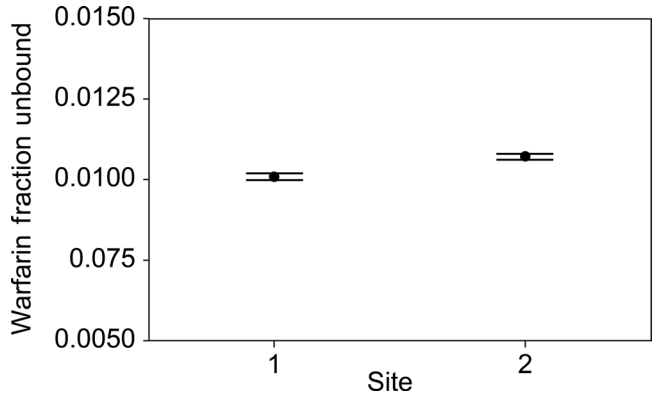

Figure 6. Grand mean values for rat warfarin PPB for two Merck sites. Error bars represent overall standard deviations. For site 1, n = 5983. For site 2, n = 250.

control chart for 16 different batch runs at two Merck sites. The mean unbound fraction for warfarin rat plasma, the most frequently requested species, was unbound fraction of 0.01. Horizontal bars on each batch date represented the standard deviation of the assay on that day against the overall grand mean unbound fraction. Each mean value was derived from at least 100 data points for that batch, with over 6000 measurements in total. Although the experiment was not in control from a statistical perspective, against the acceptance criteria for f_u = 0.01 of ±0.005 range, all data were within specification.

The overall warfarin data for the standardized assay, including 95% CIs, are shown in Figure 6 for each site. In contrast with the original results of the MSA assessment, data consistency across and within sites was extremely high. The effect of well position on variability was again assessed, with the errors from the 16 batch runs plotted by well position in the 96-well plate. Figure 7 shows a plot of error rate (z-axis) against row and column positions within the 96-well plate. No specific well had an incidence rate greater than 15% of the total number of errors. In addition, edge effects because of exterior rows and columns were not observed.

The performance of the standardized method was evaluated against the predefined acceptance criteria experiment conducted for six commercial compounds including warfarin and clozapine on 3 separate days for each site (n = 6 replicates with

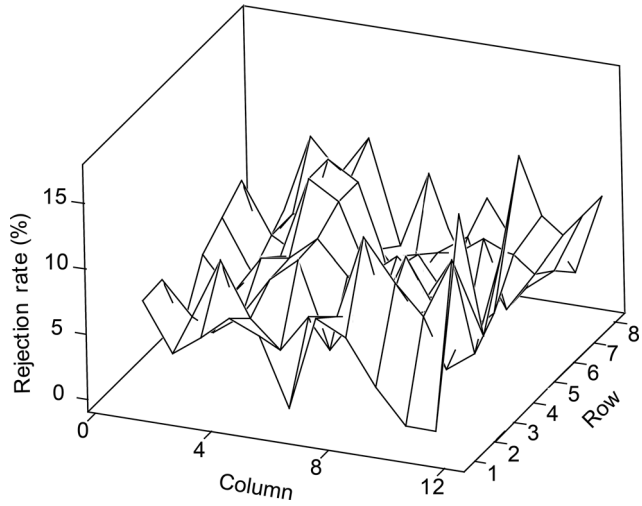

Figure 7. Rejection rate for warfarin in-well control as a function of position within 96-well plate.

warfarin in-well control). Results are shown in Table 4. Day-to-day reproducibility was acceptable for mean values from a given site, and the overall mean for each site was within the acceptable range. In addition, the majority of observed experimental values were within the acceptance criteria of literature values, with the exception of one value for verapamil and two values for diltiazem.

The failure rate as pertains to data reporting capability was also assessed across several experimental batches from day-to-day for more than 20 experiments. The proportion of compounds for which data could not be reported ranged from 0% to 13% (p = 0–0.13). The root cause of the anomalously high failure rate of 13% was subsequently corrected. In general, the failure rate was a mean of 8% not reportable, which was well within the desired range. The greater stringency provided by the established criteria was routinely met, generating greater confidence in the data generated, and improving overall data quality.

The standardized assay offers a number of advantages over the previous protocols including requiring no additional time to conduct the experiment. Changing the number of replicates from three to six has decreased the overall capacity by half, and some incremental costs including reagents, particularly plasma, have increased.

CONCLUSIONS

Critical examination of the sources of variability for the measurement of PPB using a 96-well equilibrium dialysis device have been rigorously evaluated using statistical methods including MSA and DOE. As a result of this analysis, important experimental conditions have been standardized and implemented. The most significant contributor to variability in PPB determinations was loss of physical integrity of the equilibrium dialysis membrane, most likely because of errors during pipetting. This may be because of nicks or tears from pipette tips, improper device assembly, or leakage during the incubation. The use of an in-well control provides a quick disaster check, allowing data-driven outlier rejection and consequently more rigorous results generation.

Table 4. Comparison of Reproducibility and Repeatability Across Two Merck Sites for Six Test Compounds ($n = 4$–6 for Each Measurement)

Analyte	Site 1	Site 2	Literature f_u[14,24]	Acceptable Range
Clozapine	0.0920 ± 0.0111	0.1034 ± 0.0070	0.101	0.0557–0.1443
	0.0788 ± 0.0076	0.1231 ± 0.0085		
	0.1121 ± 0.0122	0.1301 ± 0.0093		
Diltiazem	0.4327 ± 0.0517	0.5065 ± 0.0268	0.35	0.20–0.50
	0.3644 ± 0.0363	0.5401 ± 0.0127		
	0.4222 ± 0.0474	0.5720 ± 0.0206		
Fluconazole	0.8098 ± 0.1023	0.7949 ± 0.0363	0.89	0.49–1.0
	0.7922 ± 0.0326	0.8078 ± 0.0213		
	0.7830 ± 0.0415	0.8372 ± 0.0313		
Fluoxetine	0.0601 ± 0.0104	0.0656 ± 0.0070	0.06	0.044–0.076
	0.0253 ± 0.0043	0.0816 ± 0.0117		
	0.0519 ± 0.0110	0.0728 ± 0.0039		
Verapamil	0.1732 ± 0.0199	0.2473 ± 0.0067	0.16	0.052–0.278
	0.1518 ± 0.0299	0.2869 ± 0.0063		
	0.2118 ± 0.0164	0.2642 ± 0.0040		
Warfarin	0.0109 ± 0.0020	0.0129 ± 0.0008	0.01	0.007–0.013
	0.0138 ± 0.0018	0.0147 ± 0.0012		
	0.0106 ± 0.0018	0.0138 ± 0.0007		

Objective mathematical data acceptance criteria have been set for both the in-well control and the range of analyte variability, with a sliding scale tied to analyte-binding characteristics. Another key change has been to increase the number of replicates per experiment from three to six, which enables outliers to be rejected but data still reported without repeating the assay. The resulting assay rapidly generates high-quality data with a very low-failure rate.

ACKNOWLEDGMENTS

The authors would like to thank their Merck colleagues who participated in discussion and debate, including Jim Yergey, Chris Gibson, Chris Kochansky, Iain Martin, Ian Knemeyer, Randy Miller, Andreas Harsch, Kevin Bateman, Nancy Agrawal, Jane Harrelson, and Rick Morrison. We would also like to acknowledge significant contributions to the definition of the acceptance criteria by multiple former colleagues at the Organon research site in Newhouse, Scotland.

REFERENCES

1. Bateman KP, Cohen L, Emary B, Pucci V. 2013. Standardized workflows for increasing efficiency and productivity in discovery stage bioanalysis. Bioanalysis 5(14):1783–1794.
2. King LE, Leung S, Ray C. 2013. Discovery fit-for-purpose ligand-binding PK assays: What's really important? Bioanalysis 5(12):1463–1466.
3. Xu X, Ji QC, Jemal M, Gleason C, Shen JX, Stouffer B, Arnold ME. 2013. Fit-for-purpose bioanalytical cross-validation for LC–MS/MS assays in clinical studies. Bioanalysis 5(1):83–90.
4. Nilsson LB. 2013. The bioanalytical challenge of determining unbound concentration and protein binding for drugs. Bioanalysis 5(24):3033–3050.
5. Smith DA, Di L, Kerns EH. 2010. The effect of plasma protein binding on in vivo efficacy: Misconceptions in drug discovery. Nat Rev Drug Discov 9(12):929–939.
6. Bohnert T, Gan LS. 2013. Plasma protein binding: From discovery to development. J Pharm Sci 102(9):2953–2994.
7. Trainor GL. 2007. The importance of plasma protein binding in drug discovery. Expert Opin Drug Discov 2(1):51.
8. Howard ML, Hill JJ, Galluppi GR, McLean MA. 2010. Plasma protein binding in drug discovery and development. Comb Chem High Throughput Screen 13(2):170–187.
9. Banker MJ, Clark TH, Williams JA. 2003. Development and validation of a 96-well equilibrium dialysis apparatus for measuring plasma protein binding. J Pharm Sci 92(5):967–974.
10. Vuignier K, Schappler J, Veuthey JL, Carrupt PA, Martel S. 2010. Drug–protein binding: A critical review of analytical tools. Anal Bioanal Chem 398(1):53–66.
11. Huang JD. 1983. Errors in estimating the unbound fraction of drugs due to the volume shift in equilibrium dialysis. J Pharm Sci 72(11):1368–1369.
12. Tozer TN, Gambertoglio JG, Furst DE, Avery DS, Holford NH. 1983. Volume shifts and protein binding estimates using equilibrium dialysis: Application to prednisolone binding in humans. J Pharm Sci 72(12):1442–1446.
13. Kochansky CJ, McMasters DR, Lu P, Koeplinger KA, Kerr HH, Shou M, Korzekwa KR. 2008. Impact of pH on plasma protein binding in equilibrium dialysis. Mol Pharm 5(3):438–448.
14. Curran RE, Claxton CR, Hutchison L, Harradine PJ, Martin IJ, Littlewood P. 2011. Control and measurement of plasma pH in equilibrium dialysis: Influence on drug plasma protein binding. Drug Metab Dispos 39(3):551–710.
15. Nilsson LB, Schmidt S. 2001. Simultaneous determination of total and free drug plasma concentrations combined with batch-wise pH-adjustment for the free concentration determinations. J Pharm Biomed Anal 24(5):921–927.
16. Fura A, Harper TW, Zhang H, Fung L, Shyu WC. 2003. Shift in pH of biological fluids during storage and processing: Effect on bioanalysis. J Pharm Biomed Anal 32(3):513–522.
17. van Liempd S, Morrison D, Sysmans L, Nelis P, Mortishire-Smith R. 2011. Development and validation of a higher-throughput equilibrium dialysis assay for plasma protein binding. J Lab Autom 16(1):56–67.
18. Carleysmith SW, Dufton AM, Altria KD. 2009. Implementing Lean Sigma in pharmaceutical research and development: A review by practitioners. R&D Management 39(1):95–106.
19. Larsen GA. 2003. Measurement system analysis in a production environment with multiple test parameters. Qual Eng 16(2):297–306.
20. Khan RM. 2013. Design of experiments in problem solving and data analysis using Minitab: A clear and easy guide to Six Sgma methodology. Chichester, UK: John Wiley & Sons, Ltd.

21. Brook Q. 2010. Lean Six Sigma and Minitab. Opex Resources Hampshire, UK.
22. Seto S, Bateman KP, Gunter B. 2002. Development of generic liquid chromatography-mass spectrometry methods using experimental design. J Am Soc Mass Spectrom 13:2–9.
23. Junker B. 2010. Kaizen for improvement of rapid protein production for early reagent protein quantities. Biochem Eng J 49:435–444.
24. Brunton LL, Lazo JS, Parker KL. 2005. Goodman & Gilman's The Pharmacological Basis of Therapeutics. 11th ed. McGraw-Hill New York, NY.