
Research Article

Calculations for Adjusting Endogenous Biomarker Levels During Analytical Recovery Assessments for Ligand-Binding Assay Bioanalytical Method Validation

John F. Marcelletti,^{1,3} Cindy L. Evans,¹ Manju Saxena,² and Adriana E. Lopez¹

Received 16 December 2014; accepted 19 March 2015; published online 23 April 2015

Abstract. It is often necessary to adjust for detectable endogenous biomarker levels in spiked validation samples (VS) and in selectivity determinations during bioanalytical method validation for ligand-binding assays (LBA) with a matrix like normal human serum (NHS). Described herein are case studies of biomarker analyses using multiplex LBA which highlight the challenges associated with such adjustments when calculating percent analytical recovery (%AR). The LBA test methods were the Meso Scale Discovery V-PLEX® proinflammatory and cytokine panels with NHS as test matrix. The NHS matrix blank exhibited varied endogenous content of the 20 individual cytokines before spiking, ranging from undetectable to readily quantifiable. Addition and subtraction methods for adjusting endogenous cytokine levels in %AR calculations are both used in the bioanalytical field. The two methods were compared in %AR calculations following spiking and analysis of VS for cytokines having detectable endogenous levels in NHS. Calculations for %AR obtained by subtracting quantifiable endogenous biomarker concentrations from the respective total analytical VS values yielded reproducible and credible conclusions. The addition method, in contrast, yielded %AR conclusions that were frequently unreliable and discordant with values obtained with the subtraction adjustment method. It is shown that subtraction of assay signal attributable to matrix is a feasible alternative when endogenous biomarkers levels are below the limit of quantitation, but above the limit of detection. These analyses confirm that the subtraction method is preferable over that using addition to adjust for detectable endogenous biomarker levels when calculating %AR for biomarker LBA.

KEY WORDS: analytical recovery; biomarkers; ligand-binding assay; matrix effects.

INTRODUCTION

Bioanalytical method validation (BMV) of ligand-binding assays (LBA) involving biomarkers endogenously present in substantial quantities in matrix (*e.g.*, serum, plasma, *etc.*) can create challenges to accurate quantitation of that analyte as a spiked entity in the sample (1–3). This is because the assay frequently cannot distinguish between added analyte and the endogenous component. The challenge of adjusting for endogenous analyte content in matrix becomes apparent when assaying validation samples (VS) and attempting to accurately quantitate the spiked component. This same challenge presents itself during selectivity determinations when multiple individual matrix samples are tested for matrix effects by addition and quantitation of spiked analytes.

Two basic methods have been described to adjust for endogenous biomarker levels in percent analytical recovery (%AR) calculations involving analyses of such VS or selectivity samples. One method (termed addition method herein) uses a comparison of the observed VS results *vs.* the expected total concentration, the latter of which being derived by addition of the nominal value of the spike plus the observed experimental value from the unspiked matrix control. The other adjustment calculation is termed the subtraction method. Thus, the concentration of the tested biomarker observed in the matrix blank is subtracted from the amount observed in the spiked VS, and the net difference *vs.* nominal spiked value is used to calculate %AR.

Both addition (*e.g.*, 4–8) and subtraction (*e.g.*, 9–13) methods are used for endogenous analyte adjustment, although the latter is the predominant calculation in the immunoanalytical field. Frequently, it is stated that the endogenous analyte concentration had been taken into account for %AR calculations, but the exact method was ill-defined (*e.g.*, 14–16). As will be described in “Results,” the choice of calculation method can be a critical factor regarding conclusions on the reliability of the assay and resulting data. Considering that both addition and subtraction methods are

¹Immunoanalytical Operations, Tandem Labs, Laboratory Corporation of America® Holdings, 13112 Evening Creek Drive South, San Diego, California 92128, USA.

²Immunochemistry Services, Intertek Pharmaceutical Services, San Diego, California, USA.

³To whom correspondence should be addressed. (e-mail: jfmpd@yahoo.com)

currently being used, it is essential to describe in full the calculation employed for %AR determinations.

Opinions have been put forth of late supporting the subtraction method, although arguments for or against the addition method have not been widely published. The seminal article by the AAPS Ligand-Binding Assay Bioanalytical Focus Group described both addition and subtraction for calculating %AR when endogenous signal is present within the range of a LBA (17). However, the relative merits of subtraction *vs.* addition methods were not addressed in that 2003 publication. Specifics about %AR calculations were not substantially discussed in the recently released 2013 FDA BMV draft guidance, nor in the Crystal City V workshop report (18). Nevertheless, influential publications support use of the subtraction method (*e.g.*, 19–21). Moreover, recent white papers on biomarker BMV by the Global CRO Council and by the Global Bioanalysis Consortium Harmonization Team state that the subtraction method should be used for adjusting endogenous biomarker content in selectivity determinations (22,23).

Crucial to this discussion and to the generalized appreciation for the importance of calculation method are published results showing direct comparisons of the two methods using experimental data. Such analyses have not been widely published. Therefore, we undertook just such a comparison using examples from our ongoing BMV studies in normal human serum (NHS) with two Meso Scale Discovery (MSD®) multiplex cytokine assays, the Proinflammatory Panel 1 and Cytokine Panel 1 systems. Endogenous levels of the 20 cytokines were observed in NHS to varied degrees and had to be accounted for during BMV.

Case studies were selected based on the presence of detectable endogenous analytes and used to directly compare %AR results with subtraction and addition calculation methods. As shown in “Results,” it was observed that subtraction consistently yielded more convincing %AR values under conditions where addition gave unreliable conclusions. Another issue along these lines relates to endogenous analytes that are above the limit of detection (LOD), but below the limit of quantitation (BLOQ). We observed herein that subtraction of assay signal attributable to the matrix is a feasible approach to account for detectable endogenous analyte levels in %AR calculations.

MATERIALS AND METHODS

Preparation of Calibration Standard and Validation Samples

The MSD V-PLEX™ Proinflammatory Panel 1 and Cytokine Panel 1 kits for human samples were purchased for these studies (Meso Scale Discovery, Rockville, MD, catalog numbers K15049G and K15050G, respectively). The kits were supplied with lyophilized calibration standard (CS) blends, a mixture of the 10 cytokines to be measured by each respective multiplex panel. The CS blend was reconstituted with MSD assay buffer for use in standard curves and for spiking NHS as described below.

Twenty NHS samples were purchased from Bioreclamation (Hicksville, NY) and screened for suitability for use in a NHS matrix pool by running in the assays with and without spiked CS blend to monitor %AR of the individual cytokines. Sera for the

NHS pool were selected based on the lack of appreciable positive or negative matrix effects on analyses of the spike cytokines, meeting acceptance of 80–120% AR using the subtraction calculation. The NHS pool consisted of eight individuals selected from the initial lot of 20 NHS purchased. Seven levels of VS were prepared by spiking this pool of NHS with the reconstituted CS blend at escalating concentrations. All VS contained at least 95% NHS. VS and any reconstituted CS preparations were stored at -70°C . All kit components had an expiration of 3 years from the date of manufacturer, while reconstituted preparations were given a 30-day expiration.

Multiplex Assay System

The assay methods followed closely to those of the product insert, except as required to incorporate NHS matrix. Briefly, reconstituted CS blend was serially diluted in assay buffer to obtain the desired concentrations for the standard curves. Diluted calibrators were added 50 μL /well directly to the plate without further dilution. VS and NHS samples were diluted in assay buffer for a final minimum required dilution (MRD) of 1:2 (50%; Proinflammatory Panel) and 1:4 (25%; Cytokine Panel) when added 50 μL /well. All determinations were assayed in duplicate (two wells). The plate was incubated 2 h at room temperature with orbital shaking. The wells were washed with buffer (PBS with 2% Tween-20) before the addition of 25 μL /well assay diluent containing a mixture of 10 SULFO-TAG™-labeled, cytokine-specific detection antibodies. The plates were incubated again at room temperature for 2 h with orbital shaking. The wells were washed with PBS-Tween. MSD read buffer was added to the wells, and the plate loaded into a MSD SECTOR® Imager 2400 instrument where a voltage was applied to the plate electrodes, causing the capture labels to emit light as electrochemiluminescence units (ECLU). A regression of the ECLU by the standard curve samples *vs.* concentrations was performed with a 5-parameter logistic model (5-PL) using Watson LIMS™ software (Version 7.3.0.01, Thermo Fisher Scientific, Inc., Waltham, MA) and the concentrations of analytes in the samples determined following adjustment for dilution.

Relevant BMV Details

The BMV data presented herein were derived in a manner consistent with the principles of GLP referred to hereafter as regulated BMV. Lower limit of quantitation (LLOQ) is defined by the VS with the lowest passing concentration of analyte in regulated BMV (18,24,25). Less formal assay characterization protocols (variously termed exploratory, nonregulated, and method qualification) can use other means to define LLOQ as described in the “Discussion” (25–27).

Standard regulated BMV parameters were assessed with these two biomarker panels, including accuracy and precision (A&P), selectivity, stability, *etc.* Only the A&P runs will be discussed for this study. A total of six A&P runs were conducted for each multiplex panel over multiple days by two analysts. Each run consisted of a 12-point CS blend curve, 3 buffer blanks, and 3 replicate determinations of the 7 VS and NHS matrix pool, all of which assayed in duplicate. The CS

curve preparation was made fresh for each assay run, while frozen VS were used for all other determinations. Acceptance criteria as applied to these analyses included percent coefficient of variation (%CV) $\leq 20\%$ except for LLOQ where $\leq 25\%$ was applicable. Percent AR was considered acceptable at 80–120%, except for LLOQ where 75–125% was applicable.

Calculations

Percent total error (%TE) was used to describe accuracy and precision from multiples runs of VS. Percent TE as used herein is the sum of inter-assay accuracy (percent relative error, %RE) plus inter-assay precision (percent coefficient of variation, %CV). Typical acceptance criteria are TE $\leq 30\%$, except at LLOQ where $\leq 40\%$ is applicable (24). Percent spike was calculated by adding the nominal spike value with the mean endogenous concentration of the matrix blank control and dividing the nominal spike concentration by the total in the VS and multiplying by 100. LOD was defined using the mean buffer blank ECLU plus 2.5 times the standard deviation of that mean value.

Raw mass units (e.g., pg/mL) reported for spiked VS, as generated by regression analysis with Watson software, reflects the combined presence of the spiked entity plus the endogenous component. Unadjusted %AR values are derived by division of these raw mass units by the respective spike mass unit values multiplied by 100.

The addition calculation method involved division of the observed VS analytical result (measured concentration) by the expected results, which is the sum of the nominal spike plus the endogenous biomarker concentrations. The endogenous biomarker concentrations were determined by concurrent analysis of the NHS matrix blank during the A&P run of the VS. The equation for the addition %AR calculation is as follows:

% Analytical Recovery by Addition

$$= \frac{\text{Measured Concentration}}{\text{Expected Concentration}} \times 100$$

(Nominal Spike Conc. + Endogenous Conc.)

The subtraction calculation method involved elimination of the endogenous component from the %AR determination. Thus, the measured amount of endogenous analyte was subtracted from the total observed analyte concentration in the VS, which was subsequently divided by the nominal spike concentration to yield AR. The equation for the subtraction %AR calculation is as follows:

% Analytical Recovery by Subtraction

$$= \frac{\text{Net Measured Spike}}{\text{Nominal Spike Concentration}} \times 100$$

(Measured Conc. - Endogenous Conc.)

Adjustments by subtraction were also conducted using assay signal, ECLU. This adjustment involved calculating the

net ECLU attributable to the NHS matrix in the VS by subtracting buffer blank mean ECLU from the matrix blank mean ECLU to generate a net matrix blank ECLU value. Acceptance criteria for mean ECLU values were CV $\leq 25\%$ between replicate determinations and between duplicates for each of those determination. Net matrix blank ECLU was then subtracted from the mean ECLU obtained with the NHS VS. Regression analysis vs. the buffer CS curves was subsequently conducted, and the data processed without further adjustment.

RESULTS

Case Study 1: Endogenous Analyte Adjustments Using IL-12/IL-23p40 A&P Run Data

Table I shows a comparison of addition vs. subtraction methods for endogenous analyte adjustment using data from IL-12/IL-23p40 A&P assay runs. The cumulative A&P results established that the LLOQ for the assay was 17.5 pg/mL (not shown). Endogenous IL-12/IL-23p40 was detected in the matrix blanks in the range of 40.8–48 pg/mL, depending on the run (inter-assay mean 42.3 pg/mL). Using a method described by others (1,17), parallelism was confirmed for endogenous IL-12/IL-23p40 in the NHS pool over a series of five dilutions, and following adjustment for the dilution factor, a CV of 9.0% was derived between sample means. Spike IL-12/IL-23p40 was 5.83 pg/mL for all runs, approximately 12% of total IL-12/IL-23p40 analyte levels in the VS.

Three runs are shown with unadjusted %AR failing for all runs. This confirms that accurate determinations of recovery required adjustment for the endogenous component. Run 1 shows near complete recovery of the spike (96%), with passing %AR values regardless of adjustment method. Run 2 yielded only an apparent 1.8 pg/mL net detectable spike (31% of the 5.83 pg/mL added) reflected in a failing %AR using the subtraction method for adjustment. However, since such a large part of the “expected” value was due to endogenous analyte, the addition method yielded a passing %AR of 91%. Finally, run 3 shows an example of unacceptably high recovery using the subtraction method (146% of spike) which would be considered passing using the addition method for adjustment. Thus, addition and subtraction methods for adjusting endogenous analyte concentrations in %AR calculations can yield disparate AR conclusions.

Case Study 2: Adjustment for Low Relative Endogenous IL-13 Levels with High Proportion IL-13 Spike Values

Shown in Table II are inter-assay results from 6 IL-13 A&P runs using VS1, VS2, VS3, and VS4 (VS5, VS6, VS7, not shown for simplicity since unadjusted values passed criteria). These data illustrate that the disparity in conclusions given by the subtraction vs. addition adjustment methods was not restricted to biomarkers with very high endogenous analyte levels. It was observed that endogenous IL-13 was approximately 3.7 pg/mL in the matrix blank as assessed by the standard curve. Percent spike as a proportion of total IL-13 in the VS ranged from 21 to 83% using the approximate 3.7 pg/mL endogenous IL-13 value.

Table I. Case Study 1: IL-12/IL-23p40 A&P Runs Illustrate Disparity in %AR Conclusions Using Subtraction and Addition Adjustment Methods

Run	Endogenous pg/mL	Spike pg/mL	Expected pg/mL	Observed pg/mL	Net Spike pg/mL	%AR Unadjusted	%AR by Subtraction	%AR by Addition
1	48.0	5.83	53.8	53.6	5.6	919% (Fail)	96% (Pass)	99.6% (Pass)
2	40.8	5.83	46.6	42.6	1.8	731% (Fail)	31% (Fail)	91% (Pass)
3	42.0	5.83	47.8	50.5	8.5	866% (Fail)	146% (Fail)	106% (Pass)

Shown are values obtained from three A&P runs involving VS1 for IL-12/IL-23p40. Endogenous and observed IL-12/IL-23p40 values were determined by triplicate analysis of the matrix blank and VS1, respectively. Expected, net spike, and %AR values were obtained as described in “Materials and Methods.” VS with %AR values between 80 and 120% were considering passing. Inter-assay mean endogenous level was 43.2 pg/mL. Red text denotes failed VS, while blue identifies VS that passed acceptance criteria.
%AR percent analytical recovery

Unadjusted %AR failed at all VS levels shown (Table II), confirming that accurate determinations of recovery required adjustment for the endogenous component. The addition method using 3.7 pg/mL endogenous IL-13 for adjustment yielded passing conclusions at all levels, including the 1 pg/mL spiked VS1. Consequently, from such data, it would appear that the LLOQ of the assay is 1 pg/mL and the 3.7 pg/mL endogenous analyte value was a valid determination. Using this endogenous IL-13 value with the subtraction method resulted in failed VS1 and VS2, but passed at VS3 and VS4. However, the 3.7 pg/mL value should not be considered valid since it was BLOQ using the subtraction method, so the approach of using assay signal was investigated. Only VS4 passed using assay signal, identifying the LLOQ for the assay as 18 pg/mL based on these VS results.

The validity of assay signal *vs.* mass units to calculated %AR was examined more fully by comparison of representative A&P results for IL-12/IL-23p40, IL-15, and IL-16. A modest relationship ($R^2=0.6693$) was observed between the nominal spiked VS concentrations and unadjusted observed

results ($N=42$ paired values); lack of a strong relationship reflects the endogenous analytes that had not yet been accounted for in the determinations. Using the subtraction method and mass units within the range of the assays, a much stronger relationship was observed between nominal spike and derived results ($R^2=0.9650$). This was comparable to mass units calculated by subtraction of matrix signal *vs.* nominal spike ($R^2=0.9809$). There was a very strong relationship ($R^2=0.9886$) between adjusted mass units as derived by subtracting assay signal attributable to matrix *vs.* mass units derived by subtracting endogenous mass unit concentrations within the range of the respective assays. Preliminary results with individual NHS selectivity samples confirmed utility of assay signal subtraction. However, this will have to be examined in greater detail since selectivity usually involves single determinations with and without analyte and limited inter-assay comparisons.

These results demonstrate: (1) a potential pitfall using the addition method to determine LLOQ with VS, (2) potential issues associated with using mass units derived

Table II. Case Study 2: Adjustment for Low Relative Endogenous IL-13 Levels with High Proportion IL-13 Spikes Using Inter-assay Data from Six A&P Runs

VS	Mean Endogenous	% Spike in VS	Mean Observed	Adjustment	Mean %AR	%TE	Conclusion
VS1 1 pg/mL	3.7 pg/mL	21%	4.23 pg/mL	Unadjusted	423%	337%	Failed
				Addition – pg/mL	91%	26%	Passed
				Subtracted – pg/mL	57%	165%	Failed
				Subtracted - ECLU	182%	131%	Failed
VS2 3 pg/mL	3.7 pg/mL	45%	6.25 pg/mL	Unadjusted	208%	123%	Failed
				Addition – pg/mL	94%	21%	Passed
				Subtracted – pg/mL	87%	50%	Failed
				Subtracted - ECLU	148%	70%	Failed
VS3 6 pg/mL	3.7 pg/mL	62%	9.62 pg/mL	Unadjusted	160%	70%	Failed
				Addition – pg/mL	100%	10%	Passed
				Subtracted – pg/mL	99%	17%	Passed
				Subtracted - ECLU	138%	48%	Failed
VS4 18 pg/mL	3.7 pg/mL	83%	22.1 pg/mL	Unadjusted	123%	33%	Failed
				Addition – pg/mL	107%	19%	Passed
				Subtracted – pg/mL	102%	15%	Passed
				Subtracted - ECLU	118%	13%	Passed

Shown are inter-assay values obtained from six A&P runs involving VS1, VS2, VS3, and VS4 for IL-13. Inter-assay mean endogenous and observed IL-13 values were determined by triplicate analysis ($N=18$ for each from six runs) of the matrix blank and VS, respectively. Percent AR and TE values were obtained as described in “Materials and Methods.” Thirty percent TE was considered the cut point for failed *vs.* passed VS conclusions. Red text denotes failed VS, while blue identifies VS that passed acceptance criteria.

%AR percent analytical recovery, %TE percent total error, VS validation samples, ECLU electrochemiluminescence units

BLOQ for endogenous analyte adjustments, and (3) apparently successful use of assay signal for %AR calculations involving analytes that are detectable, but not quantifiable in this assay system.

Case Study 3: Comparison of Addition vs. Subtraction Adjustment Methods with High Endogenous IL-16 Levels

As shown in Table III, case study 3 confirms the disparity between addition and subtraction %AR calculations and also indicates the potential importance of cytokine spike to endogenous proportions in achieving a passing %AR. Shown are inter-assay A&P results (six runs) for IL-16, VS1 through VS7 in NHS. The nominal spike range was 5 to 494 pg/mL. Mean endogenous IL-16 was 167 pg/mL, with spiked quantities ranging from 3 to 75% of total IL-16 content in the VS. Parallelism was established for endogenous IL-16 in the NHS pool over a series of five dilutions, and following adjustment for the dilution factor, a CV of 16.1% was derived between sample means.

Adjustments for endogenous IL-16 were necessary as indicated by unadjusted VS failing %AR at all levels (Table III). The addition method yielded passing %AR at all levels, even the very lowest VS1 with 5.1 pg/mL (3%) spike. The subtraction method by mass units identified a spike of 56.5 pg/mL as the first passing VS concentration, with failed VS1, VS2,

and VS3 for %TE reasons. These same low VS were BLOQ after subtraction of the signal attributable to NHS matrix.

These results support the following conclusions, including that: (1) the addition method for endogenous analyte adjustment gave passing results regardless of spike to endogenous IL-16 proportions, (2) subtraction using mass units or assay signal yielded matching conclusions about which VS passed vs. which failed, and (3) spikes $\geq 25\%$ may be needed to consistently observe passing %AR using the subtraction method for endogenous analyte adjustment.

Requirements for VS Spike to Endogenous Biomarker Proportions to Consistently Observe Passing Analytical Recovery %TE Results

The data shown in Fig. 1 directly examines the relationships between spiked cytokine to endogenous cytokine levels in VS vs. resulting %TE by subtraction and addition endogenous analyte adjustment methods with mass units (pg/mL). TE $\leq 30\%$ was considered passing in this analysis (horizontal red lines). The proportions of spiked cytokine (panel a) and endogenous cytokine concentrations (panel b) in the various VS for analytes IFN- γ , TNF- α , IL-8, IL-12/IL-23p40, IL-15, IL-16, and VEGF were plotted vs. the resulting inter-assay %TE from six A&P runs for each analyte using the two adjustment methods.

Table III. Case Study 3: Comparison of Addition vs. Subtraction Adjustment Methods with High Endogenous IL-16 Levels

Calculation Parameters	VS1	VS2	VS3	VS4	VS5	VS6	VS7
Nominal Spiked Conc. (pg/mL)	5.1	15.2	18.8	56.5	132	395	494
Mean Endogenous (N = 18)	167 pg/mL						
Spike = % of Total IL-16 in VS	3%	8%	10%	25%	44%	70%	75%
Observed (mean pg/mL)	170	178	179	216	295	540	649
Mean net spike observed, pg/mL	4.9	13.7	15.0	52.3	131	379	488
Mean %AR, unadjusted	3360% AR Fail	1172% AR Fail	954% AR Fail	381% AR Fail	224% AR Fail	137% AR Fail	131% AR Fail
Mean %AR and %TE with addition — pg/mL	94% AR 11% TE Pass	94% AR 11% TE Pass	98% AR 7% TE Pass	97% AR 11% TE Pass	99% AR 9% TE Pass	97% AR 14% TE Pass	99% AR 11% TE Pass
Mean %AR and %TE with subtraction — pg/mL	96% AR 295% TE Fail	90% AR 99% TE Fail	79% AR 96% TE Fail	93% AR 38% TE Pass	99% AR 16% TE Pass	96% AR 19% TE Pass	99% AR 15% TE Pass
Mean %AR and %TE with subtraction — ECLU	BLOQ N/A	BLOQ N/A	BLOQ N/A	107% AR 38% TE Pass	103% AR 29% TE Pass	98% AR 28% TE Pass	105% AR 18% TE Pass

Shown are inter-assay values obtained from six A&P runs involving buffer and matrix blanks and seven VS for IL-16 (N=18 for each from six runs). Calculations were conducted as described in “Materials and Methods.” VS with inter-assay %AR values between 80 and 120% (75–125% at LLOQ) were considering passing. Percent total error (%TE) $\leq 30\%$ was considered passing, except at LLOQ where $\leq 40\%$ was acceptable. Blue text denotes VS conditions which passed acceptance criteria, while red VS failed criteria. BLOQ N/A indicates VS that could not be accurately quantitated and could not be assessed for pass or fail determinations. Notice that %AR values are mean inter-assay results and addition or subtraction calculations using mass units of mean observed or mean endogenous as shown in this table may not yield exactly the same number
%AR percent analytical recovery, %TE percent total error, ECLU electrochemiluminescence units, VS validation samples, BLOQ below the limit of quantitation

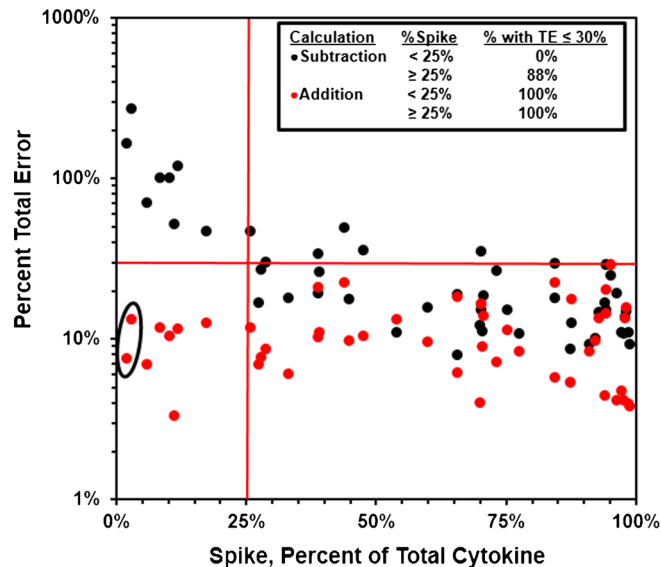
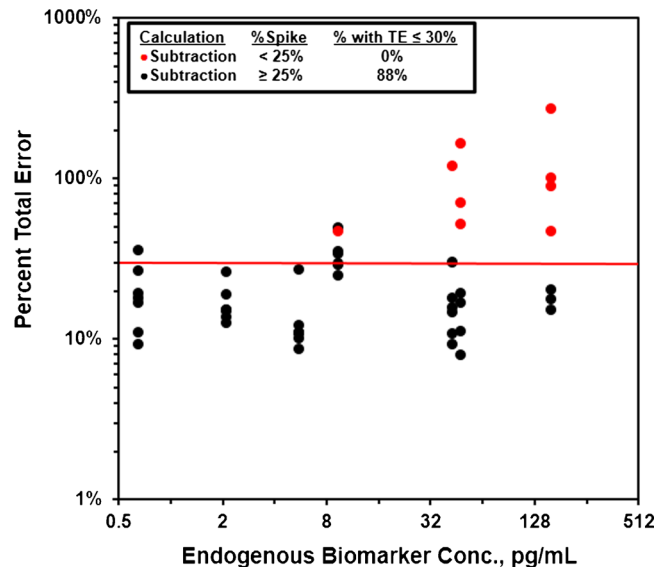
a %TE by Sub. & Add. vs. % Cytokine Spike**b %TE by Subtraction vs. Endogenous Cytokine Conc.**

Fig. 1. Successful validation samples tend to have $\geq 25\%$ spike regardless of endogenous biomarker concentrations. Cumulative A&P results involving six runs and seven VS each are shown for IFN- γ , TNF- α , IL-8, IL-12/IL-23p40, IL-15, IL-16, and VEGF. Percent TE and spike were determined as described in “Materials and Methods.” **a** presents the relationship between %TE and percent spike as a function of total respective biomarker in the VS. Red symbols for **a** represent %TE values calculated using the addition adjustment method, while black symbols represent %TE values calculated using subtraction. **b** shows %TE calculated only by subtraction vs. endogenous biomarker levels for the respective cytokines. Black symbols in **b** illustrate %TE for VS having $\geq 25\%$ spiked biomarker, and red symbols illustrate %TE for VS having $< 25\%$ spiked biomarker

As shown in panel a, using the subtraction calculation (black symbols), spikes $< 25\%$ of total tended to yield high, unacceptable %TE in the upper left quadrant of the figure. Indeed, none of the VS with $< 25\%$ spike exhibited passing %TE. Conversely, spikes $\geq 25\%$ tended to yield low %TE, as shown in the lower right quadrant of the figure, with passing %TE 88% of the time. Increasing the pass cut point for TE to 40% yielded 95% passing VS, with $\geq 25\%$ spike using subtraction. The addition calculation method yielded dramatically different conclusions, as further illustrated by the red symbols in panel a of Fig. 1. Fully 100% of the VS passed at a TE cut point of 30%, even with spikes as low as 2–3% of total cytokine (circled data points in panel a).

The relationship between baseline endogenous biomarker levels and %TE values was also investigated as shown in panel b of Fig. 1. Only %TE values calculated using the subtraction method are shown since the addition method yielded 100% passing VS regardless endogenous analyte content. Percent TE values were broken down into VS having $\geq 25\%$ spike (black symbols) vs. those with $< 25\%$ spike (red symbols). While increased variability was observed with increasing endogenous biomarker concentrations, this variability was largely restricted to VS having $< 25\%$ spike. There is an apparent relationship with increased endogenous biomarker levels which likely relates to correspondingly decreasing percent spike rather than endogenous concentrations *per se*. The data shown in Fig. 1 indicate that high A&P variability in %TE for VS is less a function of high vs. low endogenous biomarker levels in the matrix and has more to do with the proportion of spike to endogenous concentrations.

DISCUSSION

The case studies presented herein clearly support the notion that the subtraction method for adjusting endogenous biomarker content in %AR calculations produces more reliable conclusions, as contrasted to the addition method for adjustment. Three individual case studies were described in which experimental data were used to calculate %AR by the two methods. Dramatically different conclusions about acceptability of VS values were observed in each case, with results from the addition method often yielding rather unreasonable passing conclusions. Figure 1 presented the cumulative results from six A&P runs, each for seven biomarkers using seven VS (total 49 VS) for each analyte. None of the 49 VS were considered failed for %AR reasons using the addition adjustment method. This goes for VS containing as little as 2–3% spiked biomarker. This contrasts with the subtraction adjustment method where such small amounts of spiked biomarker could not be reproducibly documented with the assays at hand.

The data in Fig. 1 also indicate that successful %AR results are less a factor of very high vs. low endogenous biomarker concentrations, but rather the percent of spike to endogenous analyte content. The findings from Fig. 1 that approximately 25–30% spike to total biomarker content is needed for achieving reproducible %AR is consistent with known accuracy and precision characteristics for LBA, 30–40% TE. Thus, it is reasonable that comparable %AR conclusion might be expected when applying a 30% spike to an endogenous background of 10 or 100 pg/mL, as observed in Fig. 1, assuming sufficient sensitivity for the two respective assays.

Relationships Between the Present Findings and Other Published Works

The conclusions obtained from this study are entirely consistent with most current thinking on %AR calculations for adjusting endogenous biomarker content in spiked VS and selectivity samples (22,23). Accordingly, it is preferable to make these calculations by subtracting the endogenous component first, followed by determining the analytical recovery of the spike reference material fraction. It is nevertheless still common for the addition method to be used (4–8) and for publications to not state the calculation method employed (14–16). These observations suggest that: (1) consensus has not yet been achieved and/or (2) it is still under appreciated that the method for calculating %AR is critical to understanding the accuracy and selectivity of biomarker assays in biological matrix.

At least two factors are responsible for this lack of consensus or under appreciation. The first is that early white papers did not indicate a preference for calculation method, implying acceptability of either subtraction or addition in %AR determinations. For example, the seminal 2003 AAPS focus group white paper describes both the addition and subtraction methods, but the two calculations were not critically compared, and only the addition method was used in the examples (17). Specifics about %AR calculations were not substantially discussed in the recently released 2013 BMV draft guidance, nor the Crystal City V workshop report (18). Thus, while there are published white papers specifying the use of subtraction in selectivity assessments (22,23), the message is not being effectively delivered.

A second factor leading to common use of both calculation methods is that published data and rationale for not using addition are still lacking. The present work attempts to rectify that situation by presenting experimental data and comparative analysis of %AR calculation methods to drive conclusions on this subject. This is the first published report to our knowledge that directly compares %AR calculation methods with experimental data, providing justification for the preference of subtraction over addition in %AR determinations.

The results and conclusions described herein are specifically applicable to situations where spiked reference material and endogenous analytes exhibit a linear mathematical relationship. Situations where nonlinear contributions of these components are observed require other remedies. Physical means have been proposed to deal with matrix effects, including increasing the MRD and use of stripped or immunodepleted matrix. Matrix effects and means to address issues with endogenous biomarker levels have been discussed in detail elsewhere (1,3,20,23).

The Basis for Divergent Conclusions with the Addition and Subtraction Methods

Addition and subtraction methods for endogenous analyte adjustments will yield exactly identical numerical values only with 100% AR conditions. Deviation from 100% AR conditions, no matter how small, will always yield different %AR values between the methods. The basis for this divergence in %AR values relates to using the

endogenous component as a constant in the addition calculation, while it is eliminated with the subtraction method. The differences in %AR can be small or even negligible when using mixtures with high spike to endogenous proportions (e.g., 90 pg/mL spike with 10 pg/mL endogenous). However, even with moderate spike to endogenous proportions like 50:50, the inherent bias of the addition method can yield misleading conclusions.

To illustrate this latter point, consider the following hypothetical mixture containing 50 pg/mL spike and 50 pg/mL endogenous biomarker. The assumptions include 80–120% AR acceptance limits, and statistics like %TE and parallelism are acceptable and equal between calculations and observed results. Unadjusted assay results of 100 ± 20 pg/mL (40 pg difference) would be acceptable with addition, reflecting 80–120% of “expected” 100 pg/mL. With subtraction, unadjusted assay results would have to be 100 ± 10 pg/mL (20 pg difference) because acceptance limits concern only the spike (i.e. 80–120% of 50 pg/mL). Thus, if 80 pg/mL was the observed result, it would pass with addition (80% AR) and fail with subtraction (60% AR) because the endogenous component is treated differently between calculations.

It should be noted that the conclusions from these studies apply only to spiking defined reference materials into biological matrix samples. Addition calculations are acceptable and sometimes the only way to address %AR under certain conditions. For example, mixing of biological matrix samples containing different endogenous biomarker levels to achieve quality control samples with desired levels of that biomarker has been described and recommended (17,25,26). It is reasonable that addition would be used to derive expected concentrations under these and other circumstances.

Technical Considerations and Issues Associated with LLOQ

The distinctions between advanced biomarker method validation (i.e., regulated BMV) and protocols with less rigor (e.g., exploratory, method qualification, and category 1) have been reviewed elsewhere (1,18,20,22,25–27). It has been recommended that regulated BMV of biomarker LBA be conducted to the extent possible when the results are intended to support key claims in regulatory submissions (18). A crucial component of regulated BMV is the use of VS prepared in the same biological matrix as the authentic study samples. Multiple VS are required (generally 5, but we used 7) with spiked levels spanning the expected range of the assay. The passing VS containing the lowest level of spiked analyte defines the LLOQ of the assay with regulated BMV. This is different from biomarker analytical method qualification where the LLOQ can be defined any number of ways, including the lowest concentration of analyte that can be quantitatively determined with predefined A&P, such as spiked buffer specimens, matrix specimens containing endogenous analyte, etc.

The results reported herein can have implications for LLOQ conclusions for biomarker assays undergoing regulated BMV. A circular, and potentially fallacious, set of conclusions about passing LLOQ VS and the LLOQ of an assay could be derived using the addition, but not the subtraction, adjustment method in the context of a regulated biomarker BMV. An example of this potential problem was

observed regarding the data in Table II. As described, VS1 with 1 pg/mL IL-13 spike was the lowest passing VS in the series using the addition adjustment method. This could lead to the conclusion that 1 pg/mL is the LLOQ of the assay. Endogenous IL-13 was present in the matrix blank at ~3.7 pg/mL, an ostensibly valid conclusion since the addition method identified 1 pg/mL as LLOQ. However, two subtraction methods (*i.e.*, using mass units and ECLU) identified at least a 6 pg/mL LLOQ. These latter results cast doubt on the validity of the aforementioned 3.7 pg/mL endogenous IL-13 determination and the 1 pg/mL LLOQ conclusion. The 3.7 pg/mL was BLOQ with both subtraction methods, and the likely valid results were derived using subtraction of assay signal as described subsequently.

Adjusting for Endogenous Analyte Signal Above LOD, but BLOQ

It is conceivable that endogenous levels of a particular biomarker are normally above LOD but BLOQ, but valid data in that region are desired. For example, we were attempting in another study to verify acceptance of a 1 pg/mL LLOQ VS; the matrix blank signal was above the LOD at an approximate 0.5 pg/mL level, resulting in an approximate unadjusted 150% AR. The strategy investigated for adjusting endogenous analyte content BLOQ was to eliminate that portion of the assay signal attributable to the matrix. Thus, mean buffer blank ECLU was subtracted from the mean matrix blank ECLU to yield net matrix blank signal, which was then subtracted from all samples prepared in that matrix blank. Direct interpolation of sample data points along the buffer standard curve yielded derived mass units that required no additional adjustment. The resulting mass units were highly correlated with results obtained by subtracting mass units observed within the assay range, as described relative to case study 2. This strategy was also investigated with individual NHS selectivity samples by subtracting the net signal of the unspike sample from the spiked sample signal, with comparable conclusions. These results suggest that similar data manipulations BLOQ are valid, but more detailed studies along these lines are required to confirm that supposition.

It should be pointed out that this strategy of signal subtraction would only apply to %AR calculations during BMV and not regulated analysis of trial samples when biomarker levels fall between LOD and LLOQ. As discussed by Lee and associates (1,20), it is sometime useful to accept such sample results for exploratory biomarker analyses only, but the variability associated with working in that low analyte region should be kept in mind. We have noticed that it is common for a 10–20% CV in signal to yield >30–40% CV in derived mass units when working between LOD and LLOQ. This may explain, in part, why %TE was higher with mass unit subtraction BLOQ vs %TE with signal in Tables II and III. It is generally agreed that extrapolation BLOQ in advanced (regulated) biomarker analyses is not an accepted practice.

Assay technologies other than MSD may not be amenable to this signal manipulation strategy for adjusting BLOQ endogenous biomarker content. MSD typically has a substantial range in signal (3–4 logs) and is noted for being relatively

unaffected by matrix effects. Techniques such as ELISA, which are known to have higher biological matrix variability, may not yield such clean results and conclusions. Additional studies are therefore required to draw broad reaching conclusions about applicability of this signal subtraction strategy.

CONCLUSIONS

The presence of appreciable concentrations of biomarker(s) in matrix can complicate %AR determinations during BMV. We consider the subtraction method preferred over that of addition to adjust for detectable endogenous biomarker levels in %AR calculations, consistent with opinions of others (*e.g.*, 9–13,20,22,23). The subtraction method can use either mass units or raw assay signal, the latter of which being applicable when the endogenous analyte is present above the LOD, but BLOQ. Optimal spike to endogenous proportions are ≥25–30%, which is expected for a LBA system with %CV and %RE acceptance criteria generally 20–25% (17).

ACKNOWLEDGMENTS

This study was funded entirely by Tandem Laboratories, a subsidiary of Laboratory Corporation of America® Holdings. All authors were employed by Tandem Labs at the time of this study.

REFERENCES

1. Lee JW, Devanarayan V, Barrett YC, Weiner R, Allinson J, Fountain S, *et al.* Fit-for-purpose method development and validation for successful biomarker measurement. *Pharm Res.* 2006;23:312–28.
2. Stevenson LF. Biomarker assay development and validation for large molecules in bioanalysis: what are the challenges and how are they overcome? *Bioanalysis.* 2012;4:2657–9. doi:10.4155/bio.12.260.
3. DeSilva B, Garofolo F. Matrix interference in ligand-binding assays: challenge or solution? *Bioanalysis.* 2014;6:1029–31. doi:10.4155/bio.14.92.
4. St. Ledger K, Agee SJ, Kasaian MT, Forlow SB, Durn BL, Minyard J, *et al.* Analytical validation of a highly sensitive microparticle-based immunoassay for the quantitation of IL-13 in human serum using the Erenna® immunoassay system. *J Immunol Methods.* 2009;350:161–70. doi:10.1016/j.jim.2009.08.012.
5. Funaoka H, Dohi Y, Ohgushi H, Akahane M, Imamura T. Development of a high-specificity enzyme-linked immunosorbent assay (ELISA) system for the quantification and validation of intact rat osteocalcin. *Immunol Investig.* 2010;39:54–73. doi:10.3109/08820139.2010.534216.
6. Sand JM, Larsen L, Hogaboam C, Martinez F, Han M, Røssel Larsen M, *et al.* MMP mediated degradation of type IV collagen alpha 1 and alpha 3 chains reflects basement membrane remodeling in experimental and clinical fibrosis—validation of two novel biomarker assays. *PLoS ONE.* 2013;8:1–12. doi:10.1371/journal.pone.0084934.
7. Thomsson O, Strom-Holst B, Sjunnesson Y, Bergqvist A-S. Validation of an enzyme-linked immunosorbent assay developed for measuring cortisol concentration in human saliva and serum for its applicability to analyze cortisol in pig saliva. *Acta Vet Scand.* 2014;56:55. doi:10.1186/s13028-014-0055-1.
8. Liu QL, Yan XH, Yin XM, Situ B, Zhou HK, Lin L, *et al.* Electrochemical enzyme-linked immunosorbent assay (ELISA) for α -fetoprotein based on glucose detection with multi-enzyme-nanoparticle amplification. *Molecules.* 2013;18:12675–86. doi:10.3390/molecules181012675.

9. Bastarache JA, Koyama T, Wickersham NE, Mitchell DB, Mernaugh RL, Ware LB. Accuracy and reproducibility of a multiplex immunoassay platform: a validation study. *J Immunol Methods*. 2011;367:33–9. doi:10.1016/j.jim.2011.01.005.
10. Shukla R, Santoto J, Bender FC, Laterza OF. Quantitative determination of human interleukin 22 (IL-22) in serum using Singulex-Erenna® technology. *J Immunol Methods*. 2013;390:30–4. doi:10.1016/j.jim.2013.01.002.
11. Wu Y, Lee JW, Uy L, Abosaleem B, Gunn H, Ma M, *et al*. Tartrate-resistant acid phosphatase (TRACP 5b): a biomarker of bone resorption rate in support of drug development: modification, validation and application of the BoneTRAP kit assay. *J Pharm Biomed Anal*. 2009;49:1203–12. doi:10.1016/j.jpba.2009.03.002.
12. Kift RL, Messenger MP, Wind TC, Hepburn S, Wilson M, Thompson D, *et al*. A comparison of the analytical performance of five commercially available assays for neutrophil gelatinase-associated lipocalin using urine. *Ann Clin Biochem*. 2013;50:236–44. doi:10.1258/acb.2012.012117.
13. Bak MJ, Wewer Albrechtsen NJ, Pedersen J, Pedersen J, Knop FK, Vilsbøll T, *et al*. Specificity and sensitivity of commercially available assays for glucagon-like peptide-1 (GLP-1): implications for GLP-1 measurements in clinical studies. *Diabetes Obes Metab*. 2014;16:1155–64. doi:10.1111/dom.12352.
14. Lachno DR, Romeo MJ, Siemers ER, Vanderstichele H, Coart E, Konrad RJ, *et al*. Validation of ELISA methods for quantitation of total tau and phosphorylated-tau₁₈₁ in human cerebrospinal fluid with measurements in specimens from two Alzheimer's disease studies. *J Alzheimers Dis*. 2011;26:531–41. doi:10.3233/JAD-2011-110296.
15. Sloan JH, Siegel RW, Ivanova-Cox YT, Watson DE, Deeg MA, Konrad RJ. A novel high-sensitivity electrochemiluminescence (ECL) sandwich immunoassay for the specific quantitative measurement of plasma glucagon. *Clin Biochem*. 2012;45:1640–4. doi:10.1016/j.clinbiochem.2012.07.111.
16. Nickkova MI, Huisman H, Wynveen PM, Marc DT, Olson KL, Kellermann GH. Evaluation of a novel ELISA for serotonin: urinary serotonin as a potential biomarker for depression. *Anal Bioanal Chem*. 2012;402:1593–600. doi:10.1007/s00216-011-5583-1.
17. DeSilva B, Smith W, Weiner R, Kelley M, Smolec J, Lee B, *et al*. Recommendations for the bioanalytical method validation of ligand-binding assays to support pharmacokinetic assessments of macromolecules. *Pharm Res*. 2003;20:1885–900.
18. Booth B, Arnold ME, DeSilva B, Amaravadi L, Dudal S, Fluhler E, *et al*. Workshop report: Crystal City V—quantitative bioanalytical method validation and implementation: the 2013 Revised FDA Guidance. *AAPS J*. 2014. doi:10.1208/s12248-014-9696-2.
19. Lee J, Ma H. Specificity and selectivity evaluations of ligand binding assay of protein therapeutics against concomitant drugs and related endogenous proteins. *AAPS J*. 2007;9:E164–70.
20. Lee JW. Method validation and application of protein biomarkers: basic similarities and differences from biotherapeutics. *Bioanalysis*. 2009;1:1461–74. doi:10.4155/bio.09.130.
21. Wang J, Lee J, Burns D, Doherty D, Brunner L, Peterson M, *et al*. “Fit-for-purpose” method validation and application of a biomarker (C-terminal telopeptides of type 1 collagen) in denosumab clinical studies. *AAPS J*. 2009;11:285–394. doi:10.1208/s12248-009-9115-2.
22. Houghton R, Gouty D, Allinson J, Green R, Losauro M, Lowes S, *et al*. Recommendations on biomarker bioanalytical method validation by GCC. *Bioanalysis*. 2012;4:2439–46. doi:10.4155/bio.12.197.
23. Stevenson L, Kelley M, Gorovits B, Kingsley C, Myler H, Osterlund K, *et al*. Large molecule specific assay operations: recommendations for best practices and harmonization from the global bioanalysis consortium harmonization team. *AAPS J*. 2014;16:83–8. doi:10.1208/s12248-013-9542-y.
24. Kelley M, Beaver C, Stevenson L, Bamford R, Gegwich P, Katsuhiko Y, *et al*. Large molecule run acceptance: recommendation for best practices and harmonization from the Global Bioanalysis Consortium Harmonization Team. *AAPS J*. 2014;16:221–5. doi:10.1208/s12248-013-9553-8.
25. Khan MU, Bowsher RR, Cameron M, Devanarayan V, Keller S, King L, *et al*. Recommendations for adaptation and validation of commercial kits for biomarker quantification in drug development. *Bioanalysis*. 2015;7:229–42. doi:10.4155/bio.14.274.
26. Valentin M-A, Ma S, Zhao A, Legay F, Avrameas A. Validation of immunoassay for protein biomarkers: bioanalytical study plan implementation to support pre-clinical and clinical studies. *J Pharm Biomed Anal*. 2011;55:869–77. doi:10.1016/j.jpba.2011.03.033.
27. DeSilva B, Garofolo F, Rocci S, Martinez S, Dumont I, Landry F, *et al*. White paper on recent issues in bioanalysis and alignment of multiple guidelines. *Bioanalysis*. 2012;4:2213–26. doi:10.4155/bio.12.205.