

# Probing Liquid Chromatography–Tandem Mass Spectrometry Response Dynamics and Nonlinear Effects for Response Level Defined Calibration Strategies with Simple Methods To Expand Linear Dynamic Ranges

Rui Zhang\*

Cite This: *ACS Omega* 2024, 9, 607–617

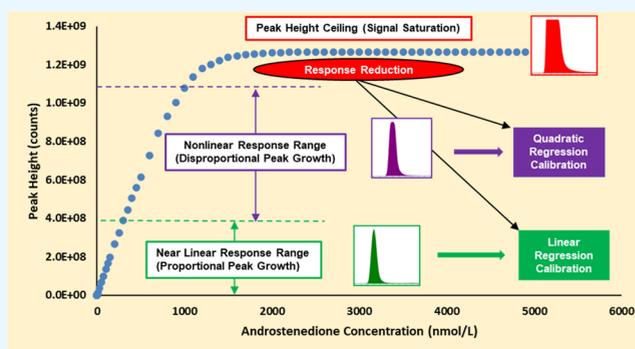
Read Online

ACCESS |

Metrics &amp; More

Article Recommendations

**ABSTRACT:** The response characteristics of liquid chromatography–tandem mass spectrometry (LC–MS/MS) serve as the basis for selecting calibration methods in quantitative analysis. LC–MS/MS inherently exhibits nonlinear detection behavior, primarily attributed to the disproportionate growth observed between peak area and peak height at elevated response levels, potentially leading to signal saturation. This disproportionate peak growth results in reduced unit response (UR), which quantifies the instrument's detection sensitivity. LC–MS/MS typically operates within a narrow near-linear response range (NLRR) due to approximately proportional peak growth, yet the NLRR width varies across different analytes or platforms. Although the inclusion of stable isotope-labeled (SIL) internal standards (IS) in LC–MS/MS analysis can mitigate certain instrument response variations, it does not eliminate the fundamental cause of nonlinearity. Moreover, the concentration range accommodated by the NLRR can significantly fluctuate at different sensitivity levels. LC–MS/MS also encounters various other nonlinear effects, including ion suppression during ionization, signal cross-contribution between the analyte/IS, and matrix effects (ME). Consequently, quadratic regression emerges as a more adaptable approach to LC–MS/MS nonlinear response dynamics, offering a broader calibration range. The application of linear regression, on the other hand, requires strict conditions. Although the signal saturation zone typically remains inaccessible to calibration methods, reducing responses by employing less-optimal selected reaction monitoring (SRM) transitions and/or lower detection gain can facilitate fitting a wide concentration range into the NLRR, thereby enabling accurate linear regression calibration. This report delves into examining the LC–MS/MS response profile, its dynamics, and major nonlinear effects through instrument response mapping to elucidate their influence on the selection of calibration methods.



## INTRODUCTION

In recent years, the application of LC–MS/MS in quantitative analysis within domains like clinical biochemistry, pharmacology/toxicology, and various other fields has surged, owing to its exceptional selectivity and specificity.<sup>1</sup> However, a discernible variability in interlaboratory results has emerged, largely stemming from the custom development of many assays without standardized or harmonized approaches.<sup>2–4</sup> Several guidelines established by government regulators and industry agencies aim to standardize the assay validation process.<sup>5–12</sup> Nevertheless, a comprehensive exploration of LC–MS/MS response characteristics, such as instrument-specific detection sensitivity profiles and linear response ranges, which might differ across instruments, remains lacking. These dynamic characteristics can significantly impact the assay performance.

LC–MS/MS response characteristics manifest as a complex interplay of numerous influencing factors. Consequently, amidst numerous reports on LC–MS/MS applications, only a scant few have delved into the response behaviors of specific LC–MS systems or across different platforms.

An earlier study attempted to scrutinize response behavior among three models of LC–MS instruments from distinct manufacturers.<sup>13</sup> They utilized the concept of “response per concentration unit” to delineate instrument sensitivity levels and

**Received:** August 21, 2023  
**Revised:** December 7, 2023  
**Accepted:** December 11, 2023  
**Published:** December 27, 2023



observed that, for all instruments studied, while the absolute response per concentration unit declined markedly at higher concentrations, the analyte/IS relative response per concentration unit, when employing a “perfect” coeluting IS, exhibited markedly diverse behaviors across these instruments. They categorized these instruments as either ionization capacity limited (ICL) systems, where ion suppression occurs before signal saturation, or detection capacity limited (DCL) systems, where detector saturation occurs before significant ion suppression. Calibration schemes, utilizing linear or quadratic calibration, were proposed based on these distinctions.

Another contemporaneous study using an ICL-type system identified the “root cause” of nonlinearity in LC–MS/MS analysis as the loss of proportionality between the analyte and IS responses.<sup>14</sup> They pinpointed a critical response level at which the instrument begins to deviate from linear response behavior. Additionally, they also demonstrated the advantages of employing a perfect IS in situations dominated by ion suppression and suggested a strategy employing two SRM transitions with different sensitivities to cover a wider linear dynamic range.

Recent advancements in LC–MS instrumentation, characterized by improved detection sensitivity, lean more toward the DCL type of systems, where signals saturate well before ion suppression takes place. Consequently, the strategies proposed in these earlier reports are likely to need revision.

This report endeavors to characterize the instrument-specific response profile by employing response mapping on two recent models of LC–MS instruments from the same manufacturer. We utilize a sample set forming a concentration gradient to achieve this aim. The focus of the study is a custom-developed and validated LC–MS/MS steroid panel assay, detecting five endogenous steroid hormones—androstenedione ( $A_4$ ), testosterone (T), 17 $\alpha$ -hydroxyprogesterone (17-OHP), dihydrotestosterone (DHT), and progesterone ( $P_4$ ). An assortment of samples containing all five analytes covering more than 5 orders of magnitude of concentration range will be analyzed to elucidate the instrument response profile. Peak height is chosen as the primary representation of the instrument response, given that the current instruments’ detection limit is determined by this parameter. The study delves into the dynamics of the instrument response profile under varying operational conditions. Additionally, the impact of mitigating signal saturation through the use of less-optimal SRM(s) and lower detector gain will be evaluated.

## MATERIALS AND METHODS

**Materials and Reagents.** The steroid standards and isotope-labeled internal standards for  $A_4$  and androstenedione- $d_5$  ( $A_4$ - $d_5$ ), T and testosterone- $d_3$  (T- $d_3$ ), 17-OHP and 17 $\alpha$ -hydroxyprogesterone- $d_8$  (17-OHP- $d_8$ ), DHT and dihydrotestosterone- $d_3$  (DHT- $d_3$ ),  $P_4$  and progesterone- $d_9$  ( $P_4$ - $d_9$ ) were procured from Cerilliant (Round Rock, Texas, USA) and Isosciences (Ambler, Pennsylvania, USA). Charcoal-stripped bovine serum was obtained from ThermoFisher Scientific (Waltham, Massachusetts, USA). LC–MS-grade water, methanol, and methyltetraabutyl ether (MTBE) were acquired from Fisher Chemical. LC–MS-grade formic acid was purchased from Lichropur (Merck KGaA, Darmstadt, Germany).

**Steroid Panel Assay: Calibration Standards and Quality Controls.** Utilizing liquid form certified reference materials (CRMs) for all steroids, a volumetric serial dilution method was employed to prepare calibration standards. The highest calibration standard (level 7) comprised 5 CRM steroids

spiked in stripped serum. Serial dilutions, one in three, were performed from the highest standard to create calibrators at levels 6–1. Level 1 concentration was validated with a one-step dilution sample to confirm the absence of propagated bias. Table 1 exhibits concentration levels for each calibration standard across all five steroids.

**Table 1. Concentration Levels of Calibration Standards for the Steroid Panel Assay**

calibration standard	$A_4$ (nmol/L)	T (nmol/L)	17-OHP (nmol/L)	DHT (nmol/L)	$P_4$ (nmol/L)
level 1	0.070	0.070	0.080	0.076	0.124
level 2	0.210	0.210	0.239	0.228	0.373
level 3	0.631	0.629	0.716	0.685	1.118
level 4	1.894	1.888	2.148	2.054	3.354
level 5	5.682	5.665	6.443	6.161	10.062
level 6	17.047	16.996	19.330	18.482	30.187
level 7	51.142	50.987	57.989	55.447	90.562

The quality control (QC) material, MassCheck steroids panel 2, encompassing all five steroids across three concentration levels, was procured from Chromsystems Instruments & Chemicals GmbH (Gräfelfing/Munich, Germany).

**Internal Standard Working Solution.** The IS working solution, containing five deuterated steroids ( $A_4$ - $d_5$ , T- $d_3$ , 17-OHP- $d_8$ , DHT- $d_3$ , and  $P_4$ - $d_9$ ), was prepared by diluting the mixed stock solution in methanol. The concentrations of the internal standards are as follows: 20.192 nmol/L for  $A_4$ - $d_5$ , 12.824 nmol/L for T- $d_3$ , 65.527 nmol/L for 17-OHP- $d_8$ , 77.261 nmol/L for DHT- $d_3$ , and 21.743 nmol/L for  $P_4$ - $d_9$ .

**Sample Preparation for Steroid Panel Assay.** The steroid panel assay employs the liquid–liquid extraction (LLE) method; 100  $\mu$ L of calibrators, quality controls, and patient samples (serum), spiked with 25  $\mu$ L of IS working solution, were extracted using 500  $\mu$ L of MTBE. After vortexing and centrifuging for 3 min, the samples were stored at  $-80$  °C for 1 h. The supernatant of each sample was transferred to a 1.5 mL Eppendorf tube, dried under vacuum, and reconstituted with 60  $\mu$ L of methanol and 40  $\mu$ L of water containing 0.1% formic acid.

**Assay Validation.** This steroid panel assay was validated according to the guidelines outlined in CLSI C57-ED1.<sup>12</sup> Parameters such as assay recovery, linearity, carryover, lower limit of detection/quantitation (LLOD/LLOQ), upper limit of quantitation (ULOQ), accuracy, intraday and interday precision, and sample stability were evaluated and met clinical requirements. The assay has been accredited by the Australian National Association of Testing Authorities (NATA) for clinical use.

**LC–MS/MS Operation Conditions.** The LC–MS/MS assay analysis and response mapping experiments were conducted using two Waters tandem mass spectrometers: Xevo TQ-S (System I) and Xevo TQ-XS (System II), both coupled with a Waters ACQUITY UPLC system (Waters Corporation, Milford, Massachusetts, USA). Experimental setups were programmed using Waters MassLynx software, and data processing was carried out using Waters TargetLynx software. Steroids underwent ionization through electrospray ionization (ESI) and were detected in the positive-ion mode utilizing the SRM method. Table 2 delineates the optimized SRM transitions for each steroid and its respective deuterated IS.

**Table 2. SRM Transitions for Five Steroids and IS**

	SRM (quantifier)	SRM (qualifier)	collision energy (eV) quantifier/ qualifier
A <sub>4</sub>	287.3 > 97.1	287.3 > 109.1	25/25
A <sub>4</sub> -d <sub>5</sub>	292.4 > 100.0		25
T	289.2 > 96.8	289.2 > 108.8	25/25
T-d <sub>3</sub>	292.3 > 97.1		22
17-OHP	331.1 > 97.0	331.1 > 109.0	28/28
17-OHP-d <sub>8</sub>	339.1 > 100.0		28
DHT	291.2 > 158.9	291.2 > 255.1	22/14
DHT-d <sub>3</sub>	294.3 > 258.2		15
P <sub>4</sub>	315.2 > 97.1	315.2 > 109.1	25/25
P <sub>4</sub> -d <sub>8</sub>	324.2 > 100.1		25

Mobile phase A (MPA) comprised LC–MS-grade water with 0.1% formic acid, while mobile phase B (MPB) constituted 100% LC–MS-grade methanol. An ACQUITY UPLC analytical column (C18 100 × 2.1 mm, 1.7 μm) facilitated the UPLC gradient, operating over an 8 min run time, as delineated in Table 3.

**Table 3. UPLC Gradients for the Steroid Panel Assay**

time	flow rate(mL/min)	MPA (%)	MPB (%)
0.0	0.4	50	50
0.2	0.4	50	50
5.0	0.4	30	70
5.2	0.4	5	95
5.9	0.4	50	50
8.0	0.4	50	50

**Concentration Gradient Sample Set for the Response Mapping Experiment.** The LC–MS/MS response profile was established using a sample set encompassing five steroids, A<sub>4</sub>, T, 17-OHP, DHT, and P<sub>4</sub>, forming a concentration gradient. The gradient initiated with a concentration of 0.01 nmol/L, followed by increments: 0.05 nmol/L, progressively rising to 0.5 nmol/L with increments of 0.05 nmol/L; then 0.5 nmol/L increments between 0.5 and 5.0 nmol/L, followed by 10.0 and 25 nmol/L, with increments of 25 nmol/L between 25 and 150 nmol/L; 50 nmol/L increments between 50 and 500 nmol/L; and 100 nmol/L increments between 500 and 5000 nmol/L.

Initial stock solutions of the five steroids in either methanol or acetonitrile were used in this experiment to ensure analyte stability and minimize matrix effects. Four stock solutions at concentrations of 5000, 500, 50, and 5 nmol/L were prepared in methanol via volumetric dilution. Various volumes of each stock were then aliquoted into Eppendorf tubes to create an incremented concentration gradient. If necessary, then any intermediate concentration could be prepared by combining equal volumes of two adjacent samples.

**LC–MS/MS Detection Sensitivity and Unit Response.** The LC–MS/MS detection sensitivity or response rate was expressed as a unit response (UR) using a molar concentration. In most cases, UR is computed from the absolute peak response, either peak area or peak height, with a unit of counts/(nmol/L). Equation 1 illustrates UR as peak height per molar concentration, where UR<sub>A</sub> represents the unit response of

analyte A, H<sub>A</sub> denotes the analyte's peak height, and C<sub>A</sub> stands for the analyte concentration in nmol/L.

$$UR_A = H_A / C_A \quad (1)$$

**Linear and Quadratic Regression Least-Squares Calibration Functions.** The linear calibration function employed in this study is described by eq 2. It utilizes the least-squares regression algorithm with 1/x weighting.

$$y = ax + b \quad (2)$$

Here, 'y' represents the response calculated from the peak area or the peak area ratio between the analyte and internal standard. The calibration curve slope is denoted by 'a', and 'x' represents the analyte concentration expressed in nmol/L. The calibration curve intercept is denoted by 'b'.

The quadratic regression least-squares calibration function in this study also employs 1/x weighting and is represented by the general formula in eq 3.

$$y = ax^2 + bx + c \quad (3)$$

Here, 'y' represents the response calculated from either the peak area or the peak area ratio between the analyte and internal standard, 'x' is the analyte concentration, 'a' and 'b' are polynomial coefficients calculated from calibration standards, and 'c' is a constant.

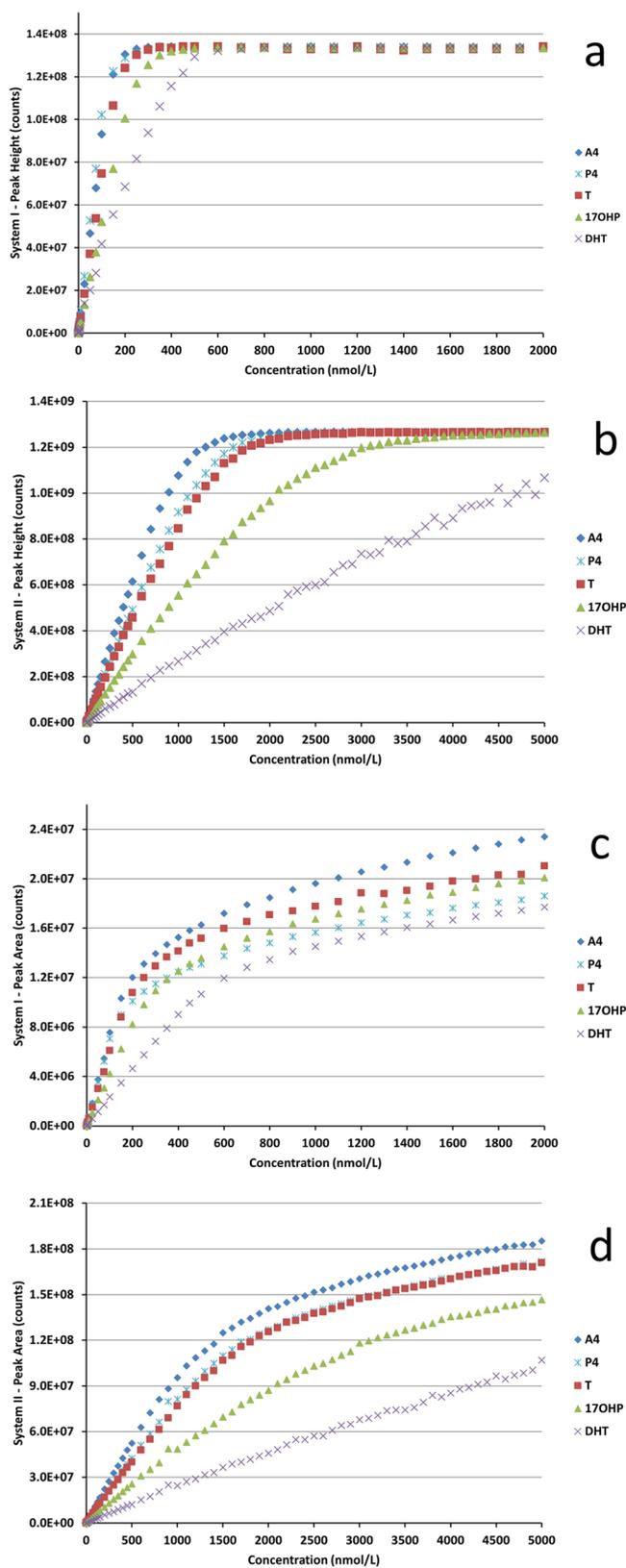
## RESULTS AND DISCUSSION

The absolute response profiles derived from response mapping for the five analytes by using two instruments are depicted in Figure 1. Graphs (a) and (b) illustrate the plots of peak height versus concentration for System I and System II, respectively. Graphs (c) and (d) display plots of peak area versus concentration for System I and System II, respectively.

A recurring pattern observed in graphs (a) and (b) is the proportional increase in peak height with concentrations up to a specific level, after which a peak height 'ceiling' is established. Notably, for DHT in System II, its peak does not reach the established peak height ceiling even at its highest concentration of 5000 nmol/L. Closer examination reveals the peak height ceiling is formed due to signal saturation which results in truncated 'flat-top' peaks. The peak height ceiling is manifested at different values for the two instruments. For System I, it approximates around  $1.34 \times 10^8$  counts, while for System II, it is approximately  $1.26 \times 10^9$  counts. An inflection zone emerges as the peak height approaches these levels. The concentrations at the inflection zone indicate that System II possesses nearly 1 order of magnitude higher linear response range compared to System I. Therefore, System II can accommodate a significantly broader concentration range within its linear response range.

While the conclusion is primarily drawn from peak height observations, similar trends are discernible when analyzing peak areas. The loss of a linear trend is evident after the establishment of the peak height ceiling. However, the behavior in peak areas differs slightly, as they can expand sideways after the peak height ceiling is reached and continue to grow. Consequently, the inflection zone is not distinctly defined with peak areas, as demonstrated in Figure 1c, d.

Ideally, LC–MS/MS, as a quantitative analytical tool, should exhibit uniform detection sensitivity across its primary or entire response range to facilitate the effective application of linear regression for calibration. The US Food and Drug Administration (FDA) guidelines for bioanalytical method validation<sup>15</sup>



**Figure 1.** Peak height vs concentration for System I (a) and System II (b). Peak area versus concentration for System I (c) and System II (d). Legends: A<sub>4</sub>, androstenedione; P<sub>4</sub>, progesterone; T, testosterone; 17OHP, 17 $\alpha$ -hydroxyprogesterone; and DHT, dihydrotestosterone.

advocate using the simplest model that adequately describes the concentration–response relationship and emphasizes justifying

the selection of weighting and the use of complex regression equations. However, these guidelines do not elaborate on how such recommendations are linked to instrument-specific response characteristics.

The LC–MS/MS detection sensitivity profile can be unveiled by examining the unit response (UR) distributions. Utilizing results from the response mapping experiments, distributions of peak height unit response (PHUR) are plotted against both concentrations and peak heights (Figure 2). Initially, there is a noticeable disparity in PHUR levels among the five analytes, reflecting differences in ionization efficiency, collision-induced dissociation (CID) efficiency, and the optimization of SRM parameters. The rapid decline in PHUR at higher concentrations, depicted in Figures 2a, b, renders the linear response range, if present, relatively narrow compared to the entire response range.

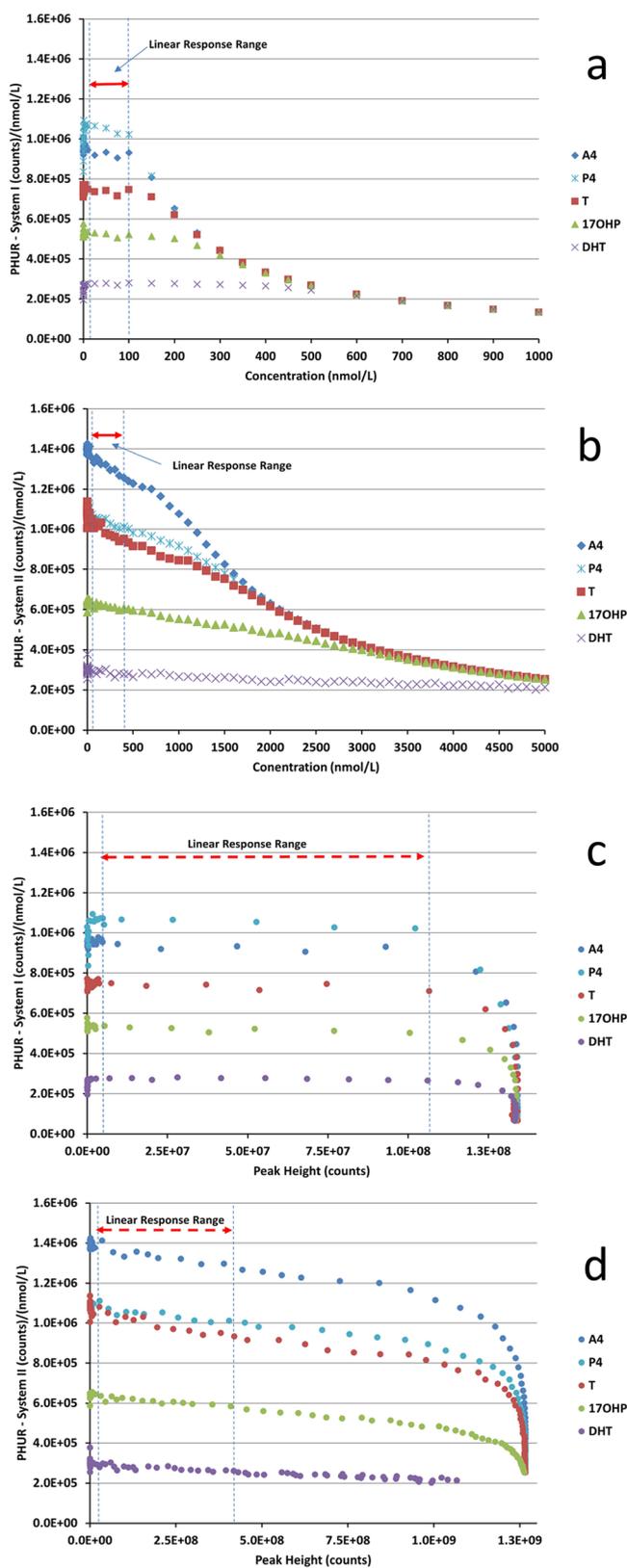
Upon scrutinizing PHUR against peak height (Figure 2c,d), it becomes evident that the instrument's response profile, in terms of detection sensitivity, can be segmented into distinct regions. At extremely low response levels, PHUR values are scattered and then transition into a relatively flat region. As the peak height approaches and exceeds the established ceiling, PHUR progressively declines. These experiments were conducted under 'ideal' conditions, utilizing only pure analytes for analysis, reflecting the predominantly nonlinear intrinsic response behavior of LC–MS/MS.

The nonlinear behavior becomes more pronounced with the increase in detection sensitivity, evident in the varying responses among different analytes from DHT to A<sub>4</sub> in Figure 2d. However, the most intriguing range is the relatively flat region, where PHUR exhibits minimal variation. This range may be aptly referred to as the near-linear response range (NLRR). NLRR, in this context, could be defined as the absolute response range of an instrument exhibiting less than  $\pm 5\%$  deviation in PHUR from the mean PHUR value calculated around the midpoint of the response range below the established peak height ceiling.

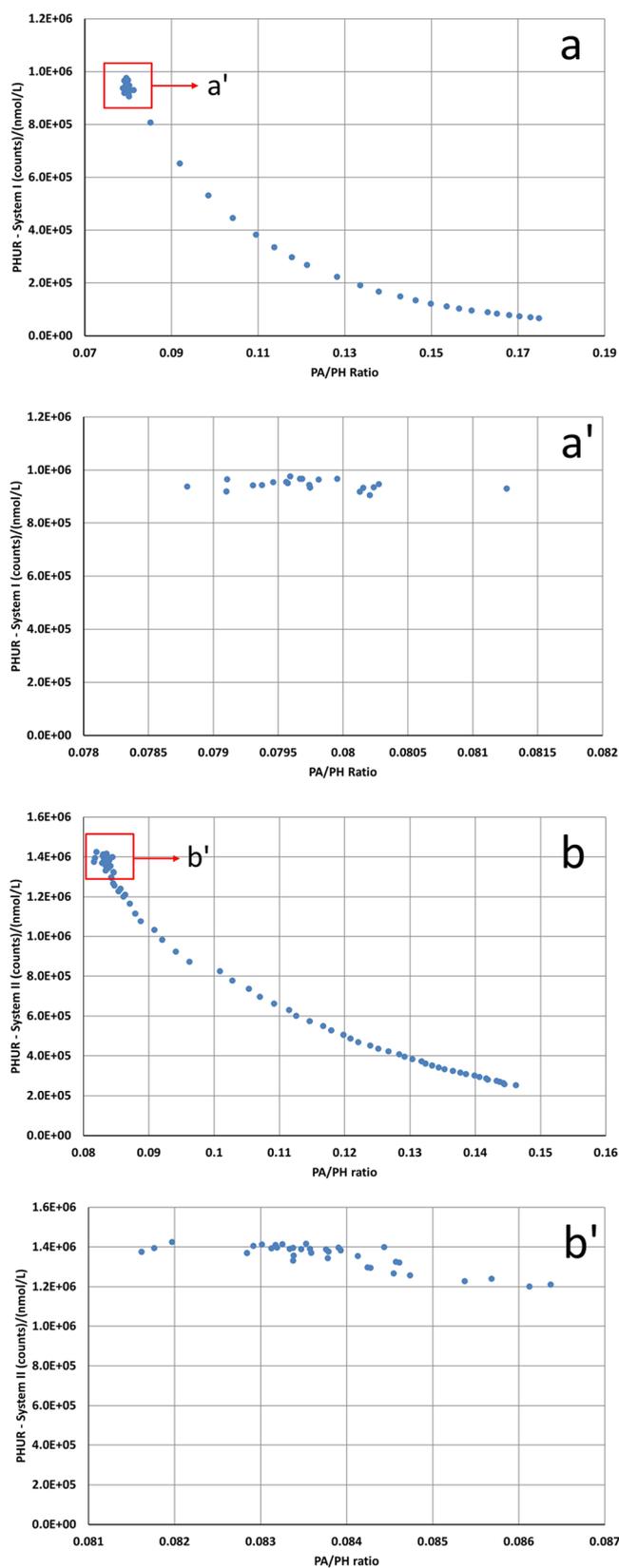
The criterion of  $\pm 5\%$  deviation is derived from the average residual errors observed in a good linear regression calibration. Notably, this criterion is specifically applied to the higher end of the response range, as at the LLOQ levels, measurement uncertainty often exceeds the  $\pm 5\%$  threshold, primarily due to random background noise levels. Across the five analytes studied in this research, the NLRR was identified on both instruments, demonstrating slight variations among them. Notably, DHT exhibited the widest NLRR with the lowest PHUR, while A<sub>4</sub> and P<sub>4</sub> showcased the narrowest NLRR with the highest PHUR on both instruments. Overall, the NLRR was observed at a peak height of approximately  $1.0 \times 10^8$  counts in System I and  $4.1 \times 10^8$  counts in System II.

The upper boundary value of the NLRR indicates that the instrument's detection linearity deviates long before the peak height reaches the established ceiling. To comprehend the underlying reasons for this phenomenon, evaluating peak quality in terms of peak shape can be insightful. One effective way to assess this is by analyzing the peak area to peak height (PA/PH) ratio, which serves as an indicator of peak proportionality. As illustrated in Figure 3, utilizing data derived from the experiments presented in Figure 1, the A<sub>4</sub> PA/PH ratio is plotted against A<sub>4</sub> PHUR for both systems.

The expanded regions from the two graphs show a relatively flat PHUR within a very narrow PA/PH ratio range, indicating consistent detection sensitivity when peaks grow proportionally



**Figure 2.** Peak height unit response (PHUR) vs concentration for System I (a) and System II (b); PHUR vs peak height for System I (c) and System II (d). Legends: A<sub>4</sub>—androstenedione, P<sub>4</sub>—progesterone, T—testosterone, 17OHP—17 $\alpha$ -hydroxyprogesterone and DHT—dihydrotestosterone.



**Figure 3.** A<sub>4</sub> peak area/peak height (PA/PH) ratio versus A<sub>4</sub> peak height unit response (PHUR) for System I (a) and System II (b); A<sub>4</sub> PA/PH ratio versus A<sub>4</sub> PHUR in the NLRR for System I (a') and System II (b').

(constant PA/PH ratio). This finding provides another perspective to delineate the NLRR. Notably, when the PA/PH ratio increases beyond this region, the peak area and peak height grow disproportionately, leading to a rapid decline in detection sensitivity (PHUR). In the absence of other nonlinear effects, this disproportional growth becomes the primary source of LC–MS/MS nonlinearity.

The underlying causes of this phenomenon could be attributed to a decreased detection efficiency of the MS detector at high response levels or a decline in the percentage of ionized samples at higher concentrations. This observation also suggests that a narrower and sharper LC peak will have a lower concentration value at the NLRR upper boundary compared to a broader and flatter LC peak. The narrower peak concentrates the sample into a shorter time frame, causing it to reach disproportional growth levels at much lower concentrations.

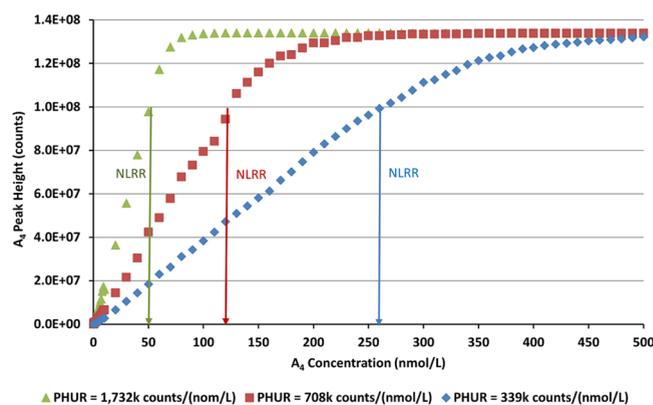
Utilizing the mean PHUR within the NLRR can serve as a quantitative measure to differentiate detection sensitivity levels among different analytes or instruments. Under identical SRM settings, System II demonstrated a significantly wider NLRR than System I across all five analytes, indicating its wider detection range in System II. Despite this, the PHUR between the two systems showed similarity, with System II exhibiting around 20–30% higher sensitivity than System I for analytes such as  $A_4$  and  $P_4$ , while DHT showed a comparable PHUR on both systems. As the two systems have almost identical ESI sources, this suggests that the wider NLRR in System II is primarily due to a better MS detector with an extended detection range rather than differences in the ionization efficiency.

The response characteristics observed in this study are not unique to the instruments used but are common across LC–MS systems, irrespective of the manufacturer and models. Earlier reports have also described similar response profiles from LC–MS instruments manufactured by different companies, emphasizing the decline in peak response per concentration unit (UR) at higher response levels across these systems.<sup>13</sup> The differences among these systems may primarily lie in the width of their NLRR for a given analyte and their peak height ceilings for signal saturation due to variations in detector capacity.

Defining an NLRR not only aids in linear regression calibration but also enables extrapolation of values beyond the calibrated concentration range up to the NLRR upper boundary. For instance, setting the concentration of the highest calibration standard for 17-OHP at 150 nmol/L with a peak height of  $4.19 \times 10^7$  counts in System II, a patient sample with 450 nmol/L of 17-OHP having a peak height of  $1.18 \times 10^8$  counts could be validated without dilution/re-extraction, as its response level is well below the NLRR upper boundary at a peak height of  $4.1 \times 10^8$  counts in System II.

In routine laboratory operations, LC–MS/MS may seem to present a consistent level of detection sensitivity. However, instrument sensitivity has the potential to vary significantly, thus altering its response profile. Various factors contribute to sensitivity fluctuation, and it can be as simple as using different solvent batches. In this study, the response mapping experiment encompassed different sensitivity levels. Figure 4 illustrates three distinct response profiles from System I selected based on their average PHUR within the NLRR using  $A_4$  as an example. The plot depicts  $A_4$  peak heights versus concentrations at three sensitivity levels.

These three conditions were derived from routine operations. The lowest sensitivity condition arose due to an extended period of continuous instrument operation in positive-ion mode. The



**Figure 4.**  $A_4$  peak heights vs concentrations at three distinct sensitivity levels of System I.

medium sensitivity level represented the instrument's average sensitivity during routine operation. Meanwhile, the high sensitivity condition emerged from using different batches of methanol for one of the mobile phases, a factor previously observed to significantly enhance sensitivity.

The average instrument sensitivity levels, measured in terms of average PHUR within the NLRR, were observed at 339k counts/(nmol/L), 708k counts/(nmol/L), and 1,732k counts/(nmol/L), respectively, in ascending order of increased sensitivity levels. Correspondingly, their concentration range within the NLRR decreased from approximately 260 to 125 and 50 nmol/L, respectively. Analysis of PHUR distributions across these three conditions demonstrated that the width of the NLRR remained unaffected by sensitivity fluctuations. However, the substantial shift in the linear concentration range emphasizes the necessity for LC–MS/MS-based assays to be designed and run at defined sensitivity levels on specific instruments.

Numerous factors, such as the addition of ionization-enhancing buffers in the mobile phase, the use of enhanced ionization methods, SRM optimization, etc., can significantly alter detection sensitivity. Consequently, for a given analyte, the linear concentration range may differ significantly. Moreover, different generations or models of LC–MS instruments from various manufacturers may exhibit different response profiles for the same analyte. Therefore, transferring an assay from one instrument to another necessitates examination of the instrument-specific NLRR at defined sensitivity levels to ensure the calibration method aligns with the desired assay concentration range.

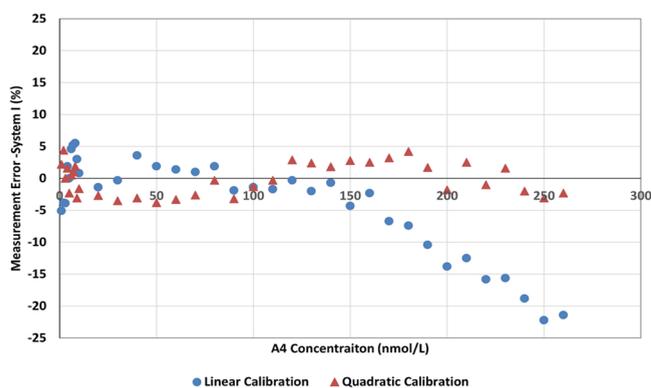
The relatively narrow response range occupied by NLRR renders a significant portion of the instrument's response range redundant when relying solely on linear calibration. Thus, it seems plausible that the region between the NLRR's upper end and the commencement of the peak height ceiling could be calibrated using the quadratic regression method. This is supported by the nonlinear distribution of PHUR in this region, as depicted in Figure 3a,b.

The adoption of quadratic regression calibration in LC–MS/MS quantitative analysis has been controversial in many cases, particularly in strictly regulated laboratories. This is due, in part, to the shifting applicable concentration range caused by LC–MS/MS response dynamics. While previous studies have attempted to develop quadratic regression calibration strategies by analyzing empirical observations of residual error distributions across different approaches,<sup>16–18</sup> these methods primarily focus on concentration ranges rather than response levels.

Consequently, their effectiveness may be compromised due to shifts in the response profile at various sensitivity levels, as depicted in Figure 4.

This study elucidates that the nonlinear response in LC–MS/MS stems from disproportional peak growth at higher response levels, as indicated by decreased PA/PH ratios. Quadratic regression calibration could better accommodate such regions, yielding more accurate results. Calculations using data from the three sensitivity conditions illustrated in Figure 4 have shown that quadratic regression least-squares calibration, particularly with  $1/x$  weighting, can cover much wider concentration ranges than linear regression calibration while maintaining good accuracy. The concentration ranges covered by quadratic regression varied from 500 to 260 and 110 nmol/L, respectively, in ascending order of increased sensitivity levels. These ranges are nearly doubled from the linear calibration results. It demonstrates that instrument response levels rather than concentration ranges or residual error distributions are the determining factors for selecting either linear or quadratic regression calibration methods.

Moreover, it is intriguing to note that with quadratic calibration the upper response levels corresponding to the highest concentrations have already entered the region of the peak height ceiling. Figure 5 illustrates the error distribution for

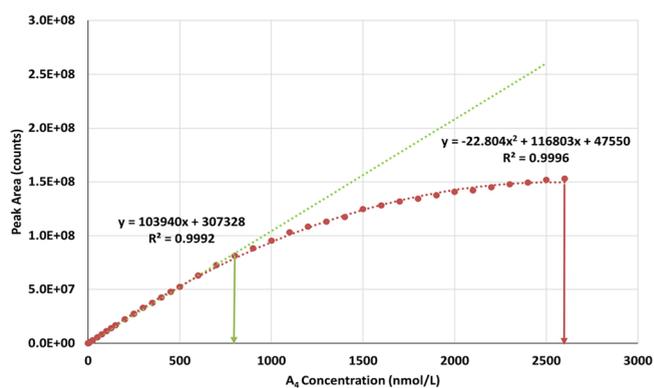


**Figure 5.** Error distributions for linear and quadratic regression calibration for  $A_4$  in System I at the medium sensitivity level shown in Figure 4.

both linear and quadratic regression for the middle sensitivity level portrayed in Figure 4. The measurement error for quadratic calibration consistently stays within the  $\pm 5\%$  range for the entire concentration range, while the error from linear calibration begins to increase above 150 nmol/L. This increase aligns with the point where the PA/PH ratio starts to decrease, signifying the onset of disproportional peak growth.

As previously discussed, System II displays a wider dynamic range compared with System I, evident in its broader concentration ranges covered by both linear and quadratic calibration methods. Illustrated in Figure 6, at an average sensitivity level in System II, linear regression least-squares calibration with  $1/x$  weighting can accommodate  $A_4$  concentrations up to 800 nmol/L. In contrast, quadratic regression least-squares calibration with  $1/x$  weighting extends the calibration concentration range up to 2600 nmol/L which also enters the peak height ceiling region.

The majority of the peak height ceiling region is typically challenging to calibrate. Attempting to use higher-order regressions in this region cannot universally yield accurate



**Figure 6.** Linear and quadratic regression least-squares calibrations with  $1/x$  weighting for  $A_4$  in System II.

results, as it involves complex conditions with multiple nonlinear effects coming into play at significant levels.

The dynamics of LC–MS/MS response profiles have revealed that while higher instrument sensitivity is a perpetual pursuit for improvement and can be advantageous for detecting low-abundance ions, it poses a double-edged sword. Higher sensitivities can notably reduce the linear dynamic range at the high concentration end, as demonstrated in Figure 4. Simultaneously, as higher sensitivity stretches concentrations within a fixed response range such as NLRR, it amplifies the variations in UR across such concentration ranges, leading to larger measurement errors when applying linear regression calibration. NLRR fundamentally lacks a true linear nature.

Conversely, when LC–MS/MS operates at lower sensitivity, its calibratable response range may accommodate a wider concentration range but its ability to detect low-abundance ions diminishes. At the lower end of the NLRR, issues extend beyond the sensitivity to include background noise levels. Random background noise is the primary cause for PHUR scattering at extremely low response ranges, as depicted in Figure 2. Higher SRM background noise can significantly reduce the analyte signal-to-noise (S/N) ratio, subsequently increasing the measurement uncertainty. Improving an analyte's measurement uncertainty may sometimes be more effective by reducing the SRM background noise level than increasing LC–MS/MS detection sensitivity. In many cases, increased instrument sensitivity tends to elevate the background noise level, trivially affecting the overall S/N ratio. Background noise is often more pronounced in the positive-ion mode than in the negative-ion mode, originating from solvents or sample matrices. Therefore, in addition to a good sample cleanup, evaluating the SRM background becomes important when employing a new batch or brand of the solvent.

Figure 7 depicts the detection of the LLOQ sample for DHT in System I at a concentration of 0.078 nmol/L using mobile-phase methanol obtained from two different suppliers. For the DHT quantifier (291.2 > 255.1), brand one exhibits a background noise level of  $1.0 \times 10^5$  counts, while brand two records  $1.5 \times 10^4$  counts. The S/N ratios for DHT quantifier peaks are 13.5 and 116.2 for brand one and brand two, respectively. The lower SRM background noise level in brand two significantly enhances the S/N ratios for detecting the DHT quantifier.

Another pertinent issue involves selecting quantifier and qualifier SRM transitions. Generally, the higher abundance transition is chosen as the quantifier. However, situations may

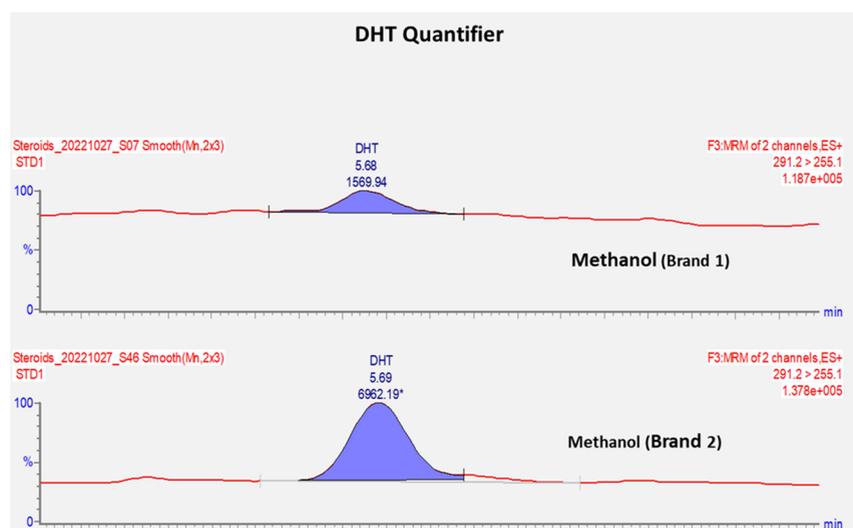


Figure 7. DHT SRM quantifier detected in System I by using a mobile phase with methanol from two different suppliers.

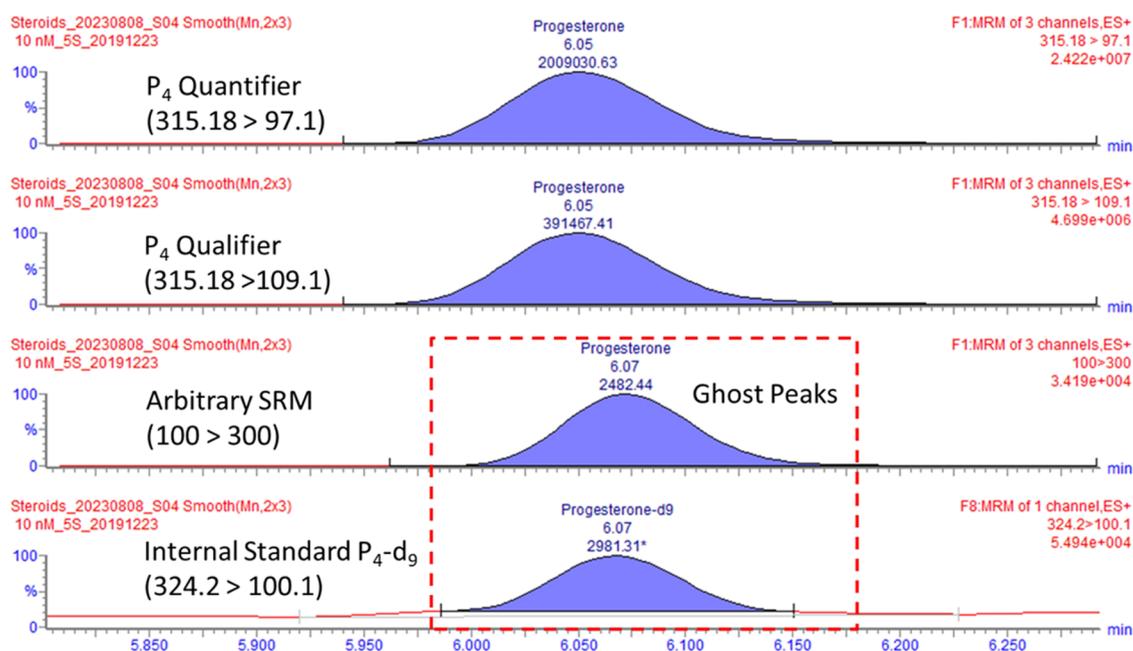


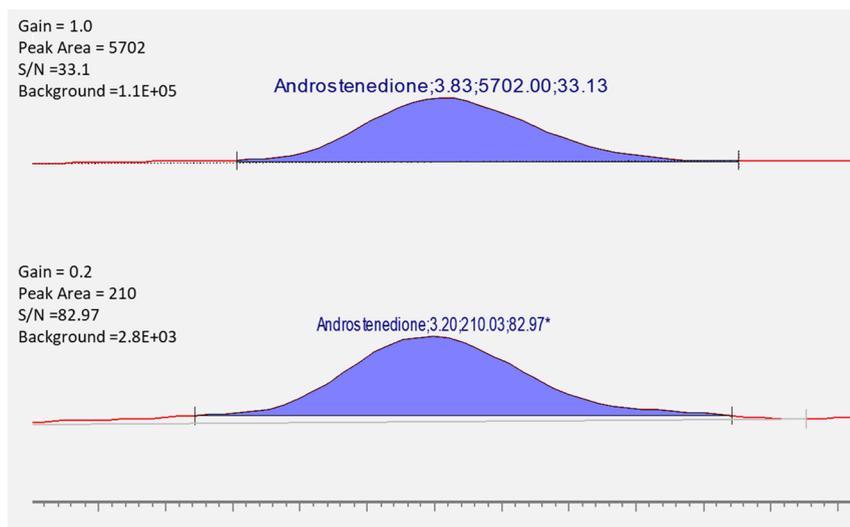
Figure 8. Instrument-generated electronic noise peak—a “ghost peak” detected in the  $P_4$ - $d_9$  SRM transition when 20 nmol/L of  $P_4$  was injected into System I.

arise where the quantifier SRM transition exhibits a significantly higher background noise than the qualifier SRM transition. In such cases, switching the quantifier and qualifier SRM transitions might enhance the S/N ratio of low-abundance ions.

The discussion so far has primarily centered on the absolute response behavior of the analyte. When utilizing stable isotope-labeled internal standards (SIL-IS), the analyte/IS relative response profile can present a more intricate scenario. While the IS can rectify variations arising from sample preparation and signal detection, its efficacy is closely tied to its molecular composition. SIL-IS labeled with isotopes such as  $^{13}\text{C}$  or  $^{15}\text{N}$  is considered superior to deuterium-labeled IS due to their closer physicochemical and biological resemblance to analytes. Moreover, their tendency to coelute with analytes may offset ion suppression and ME in the absence of signal saturation.<sup>14,19</sup> Nevertheless, as analyte and IS may contribute signals to each

other, it could introduce nonlinear effects on the relative response, particularly at higher concentration levels. According to FDA guidelines, the IS concentration should be chosen sufficiently high to suppress any analyte contributions, while its contribution to the analyte should remain under 20% of the response at the analyte's LLOQ.<sup>20</sup>

Berg and Strand presented an example of different IS types affecting the ion suppression of the analyte at high concentrations.<sup>19</sup> Their studies revealed that  $^{13}\text{C}$ -labeled IS outperformed two  $^2\text{H}$ -labeled ISs for amphetamine, compensating for ion suppression at higher concentrations and maintaining a linear correlation with assigned concentrations up to 1000  $\mu\text{mol/L}$  amphetamine concentration. The reported data were obtained using an early model of the LC-MS instrument (Waters/Premier), characterized by relatively low detection sensitivity (ICL type system). In this scenario, the peak height



**Figure 9.** Detection of 0.05 nmol/L  $A_4$  quantifier (287.25 > 97.11) with two different gain values.

ceiling was likely never reached; instead, ion suppression became the predominant effect at very high concentrations. If the same experiment had been conducted on a DCL-type instrument (such as those in the current study), the peak response might have reached the saturation level well below 20  $\mu\text{mol/L}$  (as inferred from reported data where IS began showing suppression). Consequently, the relative response between saturated peaks and coeluting IS would likely lose linearity beyond such levels. Hence, the IS's capability to rectify ion suppression significantly diminishes on these types of instruments. Moreover, as IS responses tend to be severely suppressed at high concentrations, confirming potential analyte/IS cross-contribution becomes challenging, and the observed linear behavior may not withstand sensitivity fluctuation.

Deuterium-labeled ISs are commonly used due to their cost-effectiveness and broader availability. However, their effectiveness in addressing ion suppression and ME at high concentrations varies significantly. The difference in the retention time (RT) between the analyte and deuterated IS increases with the number of deuterium atoms in the IS molecule and fluctuates with the LC gradient. In extreme cases, there might be a baseline separation between the analyte and deuterated IS, rendering the IS ineffective in correcting any ion suppression or ME.

Unlike signal saturation and ion suppression, which typically arise at high response levels, analyte/IS cross-contributions can occur at low response levels, stemming from various sources such as impurities in reference standards, isotopic interferences, or electronic sources generated by the detection of high-abundance analytes with prolonged acquisition (dwell) time. In this study, significant analyte/IS cross-contributions in the form of a 'ghost peak' were observed.

As shown in Figure 8, when 20 nmol/L pure  $P_4$  was injected into System I and both the quantifier and qualifier were detected with a dwell time over 100 ms, a peak in the  $P_4$ - $d_9$  SRM transition was observed. Notably, this peak did not stem from impurities or isotopic interferences, as it appeared consistently across various SRM transitions including any arbitrary ones such as 100 > 300  $m/z$ . Increasing interscan delay did not eliminate it. The only effective method to minimize this peak was by shortening the dwell time for the  $P_4$  quantifier and qualifier. Additionally, it vanished when the  $P_4$  quantifier and qualifier

SRM were not detected. The RT of the actual  $P_4$ - $d_9$  peak preceded that of the  $P_4$  peaks, while the ghost peak's RT slightly lagged behind the  $P_4$  peaks, resembling an "echo" from the  $P_4$  peak detection. Its peak area increased proportionally with both  $P_4$  concentration and dwell time, yet the lowest concentration for this peak to manifest differed significantly between System I and System II, with the threshold in System II being over a hundred times higher than in System I, thereby impacting System II to a much lesser extent.

For assays utilizing a DCL-type instrument aiming to cover a wide concentration range, method design plays a pivotal role in expanding its dynamic range, especially when linear regression calibration is desired. Strategies such as optimizing SRM transitions for low-abundance ions and employing less-optimal SRM transitions (lower collision energies/shorter dwell times) for a broader dynamic range or adjusting the dilution factor between raw and processed samples, along with injecting smaller sample volumes, are viable options. However, these methods possess limitations; for instance, dilution of dry blood spot samples may pose challenges, and some analytes may lack multiple available SRM transitions.

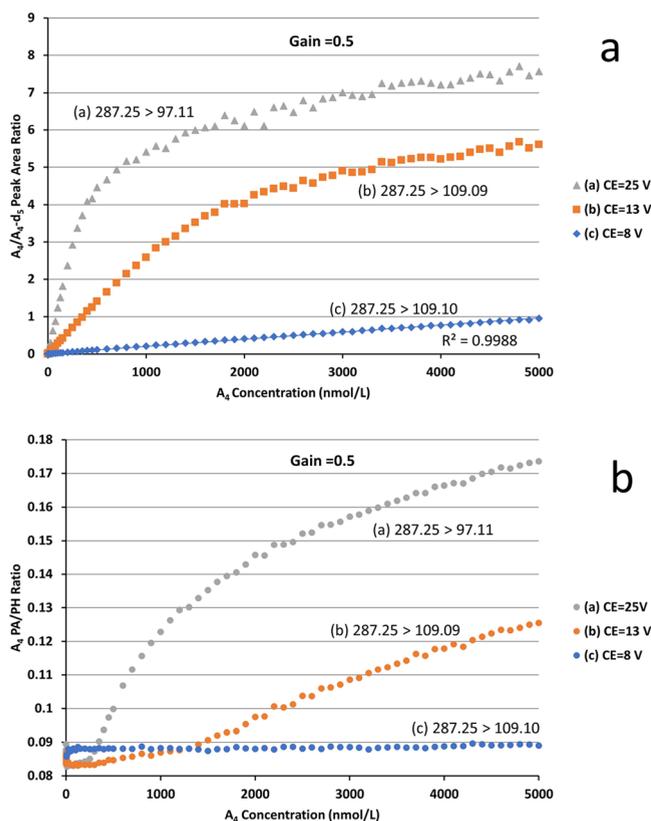
The key issue in expanding the linear dynamic range is to reduce the responses of the targeted concentration range to fit within the NLRR. In addition to the aforementioned methods, an effective approach involves lowering the LC–MS detector gain, widely applicable to most systems. Determining appropriate gain values for a desired concentration range involves analyzing the highest concentration sample at different gain values, with the selected gain value resulting in the peak height falling below the upper boundary of the NLRR.

Lowering the gain not only diminishes signal responses but also reduces background noise levels. Consequently, its impact on LLOQ level samples might be inconsequential, occasionally leading to an improved S/N ratio with a lower gain setting. For example, Figure 9 demonstrates the detection of an LLOQ level  $A_4$  sample at a concentration of 0.05 nmol/L using two different gain settings, 1.0 and 0.2, respectively. At the lower gain, a significantly smaller  $A_4$  peak is detected, yet both the background noise level and the S/N ratio have substantially improved.

Contrasted with employing two SRM transitions to cover a broader dynamic range, reducing the gain value offers a universal

advantage without specific adjustments of the SRM transitions. This aspect proves especially beneficial in multiple analyte assays covering broad dynamic ranges, where sometimes tens or even hundreds of analytes are simultaneously detected.

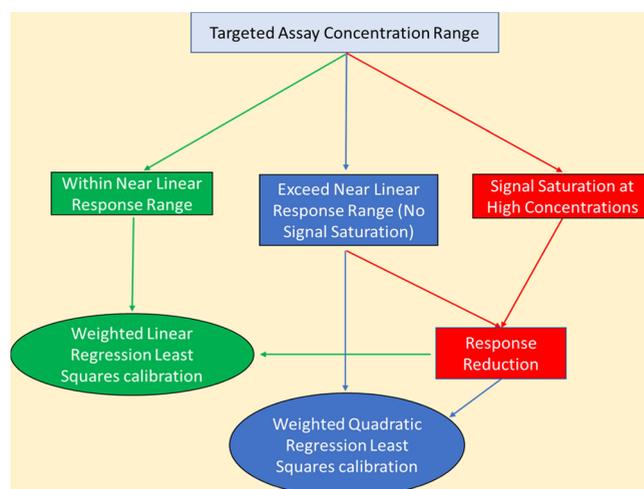
LC-MS/MS detector gain has a lower limit for each instrument, and for instance, in System I, the default lowest gain value is 0.05. However, even at the lowest gain setting, response mapping experiments for analytes such as  $A_4$  and  $P_4$  do not achieve a full linear response for the whole concentration range, encountering signal saturation at high concentrations. Therefore, combining lower gain with lower collision energy (CE) SRM transitions becomes the preferred method to attain a linear response range. Figure 10 displays three  $A_4$  SRM transitions with different CEs at a detector gain setting of 0.5 simultaneously run in System I for the response mapping experiment.



**Figure 10.** (a)  $A_4/A_4-d_5$  peak area ratio versus  $A_4$  concentration for three  $A_4$  SRM transitions. (b)  $A_4$  PA/PH ratio versus  $A_4$  concentration for three  $A_4$  SRM transitions.

The relative response of the three SRM transitions reveals that while the high CE  $A_4$  SRM still produces a response profile resembling the one depicted in Figure 1c, the  $A_4$  SRM transition with a CE of 8 V and gain value of 0.5 has achieved a linear response across the entire concentration range (Figure 10a). Similar response profiles are also observed for the other four analytes. Notably, the peak heights of the highest concentration samples (5000 nmol/L) all remain well below  $1.0 \times 10^8$  counts, which marks the upper boundary of the NLRR for System I. The linearity observed is sustained by a consistent PA/PH ratio across the entire concentration range, as depicted in Figure 10b, suggesting that this linearity is a result of proportional peak growth from this SRM transition.

Based on the observations and analysis in this study, calibration strategies can be represented in a simple plot illustrated in Figure 11.



**Figure 11.** LC-MS/MS calibration strategies are based on instrument response levels.

Prior to implementing the calibration scheme, it is advisable to characterize the instrument's response profile by conducting response mapping. If the responses of the targeted assay concentration range align well with a characterized NLRR on a particular instrument within a defined sensitivity range, then the application of linear regression least-squares calibration with specific weighting can be notably effective. Regular monitoring of instrument sensitivity is recommended to ensure that the response of the highest concentration stays within the NLRR.

When high-concentration samples produce responses that exceed the NLRR but generally do not indicate signal saturation within the high-concentration range, two options exist to address this condition. It can either be calibrated through quadratic regression calibration or utilize a response reduction method, such as using different CE SRMs or lower gain, to achieve linear calibration.

If signal saturation occurs for a significant portion of the targeted concentration range, it will necessitate response reduction before the application of either linear or quadratic calibration methods. This can be achieved through a single or a combination of several response-reduction method(s), as discussed above.

## CONCLUSIONS

The investigation into the LC-MS/MS response profile in the SRM mode via response mapping experiments has illuminated the intrinsic nonlinear nature of LC-MS/MS detection, characterized by diminishing detection sensitivity at higher response levels. The establishment of a near-linear response range within the lower response realm has facilitated the application of linear regression calibration. Additionally, quadratic regression calibration demonstrated applicability beyond the NLRR and up to the region of signal saturation.

The dynamics observed in the LC-MS/MS response profiles, influenced by variations in instrument sensitivity, emphasize the criticality of monitoring instrument sensitivity levels during routine laboratory operations. This monitoring ensures that the targeted assay concentration range is accurately measured by

using the specific calibration method employed. Notably, employing response reduction techniques, such as a lower CE SRM coupled with reduced detector gain, has proven to be a straightforward yet effective approach to achieving extended linear dynamic ranges.

## AUTHOR INFORMATION

### Corresponding Author

Rui Zhang – Department of Special Chemistry, PathWest Laboratory Medicine, Sir Charles Gardiner Hospital, Nedlands, Western Australia 6009, Australia; [orcid.org/0000-0002-2878-3348](https://orcid.org/0000-0002-2878-3348); Email: [rui.zhang@health.wa.gov.au](mailto:rui.zhang@health.wa.gov.au)

Complete contact information is available at:  
<https://pubs.acs.org/10.1021/acsomega.3c06190>

### Notes

The author declares no competing financial interest.

## ACKNOWLEDGMENTS

The author sincerely thanks Mr. Zak Glenister from PathWest Laboratory Medicine for graciously proofreading the manuscript.

## REFERENCES

- (1) Thomas, S. N.; French, D.; Jannetto, P. J.; Rappold, B. A.; Clarke, W. A. Liquid chromatography–tandem mass spectrometry for clinical diagnostics. *Nat. Rev. Methods Primers* **2022**, *2*, 96 DOI: [10.1038/s43586-022-00175-x](https://doi.org/10.1038/s43586-022-00175-x).
- (2) Vogeser, M.; Sager, C. Pitfalls associated with the use of liquid chromatography–tandem mass spectrometry in the clinical laboratory. *Clin Chem* **2010**, *56*, 1234–1244.
- (3) Greaves, R. A guide to harmonization and standardization of measurands determined by liquid chromatography–tandem mass spectrometry in routine clinical biochemistry. *Clin. Biochem. Rev.* **2012**, *33*, 123–132.
- (4) Li, S. J. Standardization of LC-MS/MS in a clinical laboratory. *J. Chromatogr. Sep. Tech.* **2015**, *6* (2), No. e128.
- (5) U.S. Department of Health and Human Services Food and Drug Administration. *Guidance for Industry: Bioanalytical Method Validation, Draft Guidance*; U.S. Department of Health and Human Services Food and Drug Administration, 2013.
- (6) AOAC Guidelines for Single-Laboratory Validation of Chemical Methods for Dietary Supplements and Botanicals. *Official Methods of Analysis*, 19th edition; AOAC International: Gaithersburg, MD, 2012; Appendix K.
- (7) SANTE/11945/2015 (Until 01.01.2016: SANCO/12571/2013). *Guidance document on analytical quality control and method validation procedures for pesticides residues analysis in food and feed*; European Commission, 2015.
- (8) European Commission Decision 2002/657/EC implementing Council Directive 96/23/EC concerning the performance of analytical methods and the interpretation of results *Off. J. Eur. Commun.* **2002**, *8*(36), L221.
- (9) *Guidance on bioanalytical method validation*; European Medicines Agency, 2011.
- (10) Pierson-Perry, J. F. *Protocols for Determination of limits of detection and limits of quantitation; approved guideline*, 2nd edition EP17A2E (R2018); Clinical and Laboratory Standards Institute, 2012.
- (11) CLSI. *Liquid Chromatography-Mass Spectrometry Methods; Approved Guideline. CLSI document C62-A*; Clinical and Laboratory Standards Institute: Wayne, PA; 2014.
- (12) *Mass Spectrometry for Androgen and Estrogen Measurements in Serum*; CLSI C57-ED1, Clinical and Laboratory Standards Institute, 2015.
- (13) Nilsson, L. B.; Skansen, P. Investigation of absolute and relative response for three different liquid chromatography/tandem mass spectrometry systems; The impact of ionization and detection saturation. *Rapid Commun. Mass Spectrom.* **2012**, *26* (12), 1399–1406.
- (14) Yuan, L.; Zhang, D.; Jemal, M.; Aubry, A. F. Systematic evaluation of the root cause of non-linearity in liquid chromatography/tandem mass spectrometry bioanalytical assays and strategy to predict and extend the linear standard curve range. *Rapid Commun. Mass Spectrom.* **2012**, *26*, 1465–1474.
- (15) Guidance for Industry, Bioanalytical Method Validation, US Food and Drug Administration, 2001. Available: [www.fda.gov](http://www.fda.gov).
- (16) Gu, H.; Liu, G.; Wang, J.; Aubry, A. F.; Arnold, M. E. Selecting the Correct Weighting Factors for Linear and Quadratic Calibration Curves with Least-Squares Regression Algorithm in Bioanalytical LC-MS/MS Assays and Impacts of Using Incorrect Weighting Factors on Curve Stability, Data Quality, and Assay Performance. *Anal. Chem.* **2014**, *86* (18), 8959–8966.
- (17) Fu, Y.; Li, W.; Flarakos, J. Recommendations and best practices for calibration curves in quantitative LC–MS bioanalysis. *Bioanalysis* **2019**, *11* (15), 1375–1377.
- (18) Cheng, W. L.; Markus, C.; Lim, C. Y.; Tan, R. Z.; Sethi, S. K.; Loh, T. P.; IFCC Working Group on Method Evaluation Protocols. Calibration Practices in Clinical Mass Spectrometry: Review and Recommendations. *Ann. Lab. Med.* **2023**, *43* (1), 5–18.
- (19) Berg, T.; Strand, D. H. <sup>13</sup>C labelled internal standards—a solution to minimize ion suppression effects in liquid chromatography–tandem mass spectrometry analyses of drugs in biological samples? *J. Chromatogr. A* **2011**, *1218* (52), 9366–74.
- (20) US FDA. ICH Guideline M10 on Bioanalytical Method.