Very complex internal standard response variation in LC-MS/MS bioanalysis: Root cause analysis and impact assessment

Van De Merbel, Nico C; Koster, Remco A; Ohnmacht, Corey

Published in:
Bioanalysis

DOI:
10.4155/bio-2019-0122

Publication date:
2019

Special Report
SPECIAL FOCUS ISSUE| Internal Standard Variability

Bioanalysis

Very complex internal standard response variation in LC–MS/MS bioanalysis: root cause analysis and impact assessment

Nico C van de Merbel*[,1,2], Remco A Koster[1] & Corey Ohnmacht[3]
[1]Bioanalytical Laboratory, PRA Health Sciences, Amerikaweg 18, 9407 TK, Assen, The Netherlands
[2]Department of Analytical Biochemistry, Groningen Research Institute of Pharmacy, University of Groningen, Antonius Deusinglaan 1, 9713 AV, Groningen, The Netherlands
[3]Bioanalytical Laboratory, PRA Health Sciences, 10836 Strang Line Road, Lenexa, KA 66221, USA
*Author for correspondence: MerbelNicovande@prahs.com

Internal standards (ISs) are essential for the development and use of reliable quantitative bioanalytical LC–MS/MS methods, because they correct for fluctuations in the analytical response that are caused by variations in experimental conditions. Sample-to-sample differences in the IS response are thus to be expected, but a large variability often is an indication of nonoptimal sample handling or analysis settings. This paper discusses a number of cases of very complex variation of IS responses that could be attributed to analytical problems such as injection errors and sample inhomogeneity, and matrix-related issues such as degradation and increased ionization efficiency. A decision tree is proposed to help find the underlying root cause for extreme IS variability.

First draft submitted: 21 May 2019; Accepted for publication: 9 September 2019; Published online: 30 September 2019

Keywords: internal standard • LC–MS • matrix effects • response variability • root cause analysis

In LC–MS/MS bioanalysis, internal standards (IS) are used to compensate for variability in the analytical response, which originates from fluctuations in experimental conditions during sample processing and further analysis [1,2]. By selecting an IS with physicochemical properties as similar as possible to those of the analyte of interest, and adding the IS to the sample as early as possible in the analytical workflow, an optimal response normalization is obtained. For this reason, a stable-isotope labeled (SIL) form of the analyte, in particular ^{13}C- or ^{15}N-labeled, is preferred as IS for LC–MS/MS [3,4], although the use of other, less ideal molecules such as chemical analogs is sometimes unavoidable. Because of the sample-to-sample differences in, for example, composition of the biological matrix, extraction recovery, transfer and injection volumes, chromatographic performance and ionization efficiency, some variation in IS response across an analytical run can be expected and is no reason for concern; it is indeed the reason why an IS is used in the first place. However, a large variability in the IS response for some or all samples and/or a systematic difference between groups of samples often indicates that the analytical process is not under control, and in such a case investigation of the underlying root cause and evaluation of the impact on the reliability of the analytical results is needed. IS responses can vary in many different ways and it is difficult to define when exactly the variability is acceptable and when it becomes a matter of concern, but in our view a coefficient of variation (CV) of above 30% can be considered large, while above 50% the variability is extreme.

Monitoring the IS responses across an analytical run and investigating the reliability of possibly affected results is, therefore, good bioanalytical practice, and a regulatory expectation as well [5–7]. For a proper assessment of the impact of extreme IS response variability, finding the cause of the effect is desirable, and by a structured, step-by-step evaluation of available data it is often possible to do so. An important aspect to investigate is whether the IS response is only affected for study samples or also for spiked calibration and quality control samples. If the former, it is not unlikely that the biological matrix and the potential variation in its composition are responsible for the varying IS response, and that the analytical method does not play a role; in other words: that there is a pre-analytical root cause. In contrast, if all samples are affected, the observed variability must be related to the method and the root cause very

newlands press

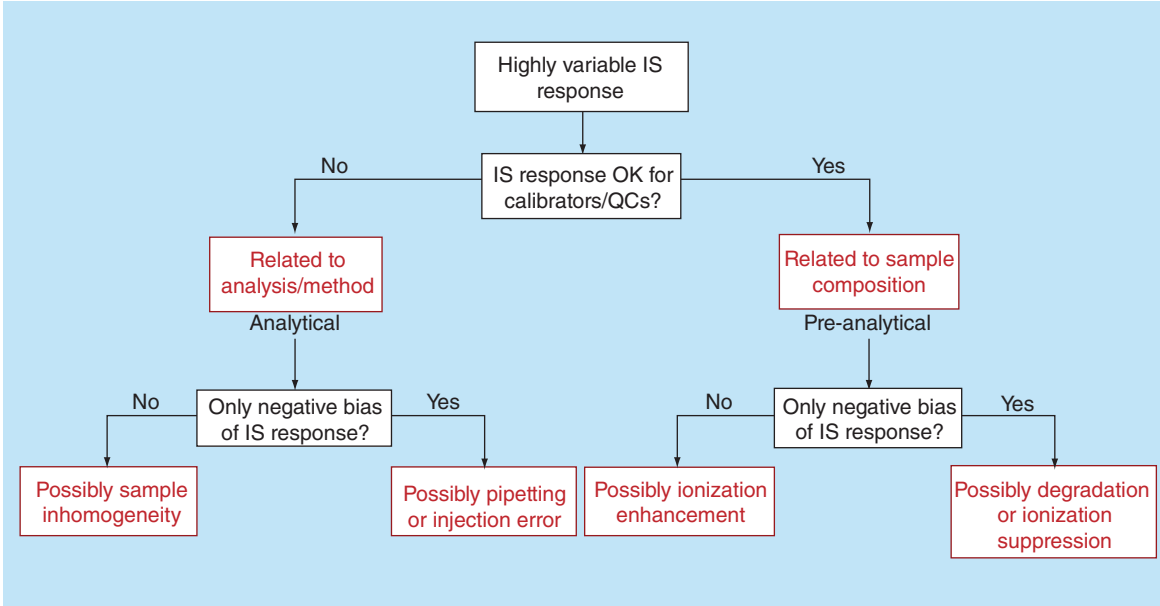

Figure 1. Decision tree to help identify the root cause for very complex variability of the internal standard response in LC–MS/MS assays.
IS: Internal standard; QC: Quality control sample.

likely is analytical. In addition, the exact way in which the IS varies can be helpful in determining the root cause. If the response only decreases compared with its normal or expected level, issues such as degradation, adsorption, a reduced injection volume and ionization suppression may be involved. If the IS response only increases, the cause may be solvent evaporation, an increased injection volume or ionization enhancement. If the effect is completely random, sample homogeneity may be compromised or multiple effects may play a role. The decision tree depicted in Figure 1 summarizes some of these issues. It is by no means exhaustive, but it does show four major general areas of IS variability and may be helpful for root cause analysis when extreme variation in IS responses is observed, as will be illustrated below for a number of case studies.

When it comes to assessment of the impact of extreme IS response variability, it has to be found out whether the analyte response varies to the same extent as that of the IS, and whether sufficient correction is thus still provided. If the IS response variability is similar for study samples and spiked samples, and acceptable accuracy is found for the spiked samples, it can be concluded that the IS is doing its job and that results are likely to be accurate. If the IS response only varies for the samples with unknown concentrations, the evaluation is more difficult and additional tests may be necessary. For example, a selection of samples with a deviating IS response could be diluted with unaffected control matrix and re-analyzed. If the IS responses of these diluted samples have come back to normal and the reanalysis concentrations are the same as the original ones, the deviating IS response in the original analysis apparently did not affect the analytical outcome.

To illustrate the relevance of IS response variability in quantitative bioanalysis, this paper describes four selected cases, from the scientific literature and our own laboratory, with an extreme variation of the IS response. Root cause analysis and impact assessment are described and discussed for all cases, and show that IS responses can be affected by very different analytical and sample-related issues.

Case 1: injection error

In an interesting paper by Tan *et al.* [8], a number of cases of unusual IS response variation were described and evaluated. One of the cases showed a complex variability of IS response across an LC–MS/MS run for clopidogrel carboxylic acid in plasma, with a random drop of up to a factor two in several samples in the later part of the run, as shown in Figure 2. Since a spiked quality control (QC) sample was affected next to a variety of study samples, it was concluded that an analytical issue rather than a sample-related effect had to be the cause of the observation. The accuracy of the affected QC sample was acceptable, and therefore it was concluded that analyte and IS had to be influenced to the same extent, at the very least in this QC sample but probably also in the other, unknown

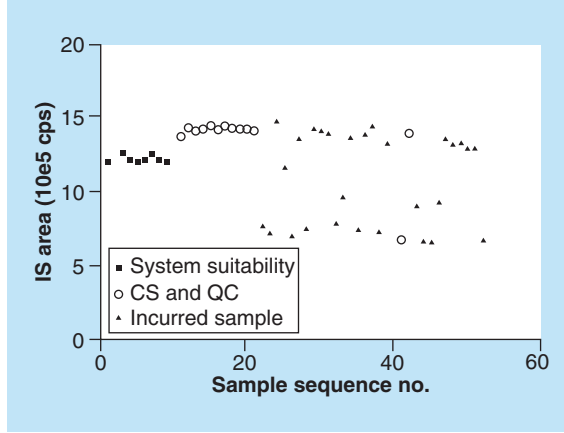

Figure 2. Response obtained for the internal standard of clopidogrel carboxylic acid in an LC–MS/MS method.
CS: Calibration sample; IS: Internal standard; QC: Quality control sample.
Reproduced with permission from [8], © 2009 Elsevier B.V. (2009).

samples in the run. The fact that the IS responses were only seen to decrease led the authors to the conclusion that an injection problem could be responsible (see also the decision tree in Figure 1).

Indeed, the partial blockage of the autosampler needle was identified as the root cause. This needle had to pierce the caps of the sample vials in the autosampler prior to aspiring the sample, which caused the formation of debris that subsequently blocked the needle to some extent. Apparently, this blockage only occurred in the later part of the run, when a relatively large number of caps had been pierced. As a result, a smaller sample volume was introduced into the LC–MS/MS system for some samples, and this resulted in reduced IS and analyte responses. The autosampler needle flush procedure appeared to be able to flush out all or part of the debris during the rest of the analysis, which explains that the IS response was restored to its normal level in subsequent injections. Clearly, it is impossible to predict if and when the needle will be blocked and if and when the blockage will be resolved, so a random IS variability was obtained. Although they do not seem to have actually tested this, the authors suggested that reinjection of the run with a different autosampler or on the same system after replacing the autosampler needle would have to result in a normal IS response plot.

When it comes to impact assessment, the situation is quite clear: the injected amount of analyte and IS is reduced to exactly the same extent, so their concentration ratio is unaffected and the obtained results are reliable, as long as the responses do not drop below the values at the normal LLOQ level.

Case 2: matrix-induced degradation

An LC–MS/MS method for the antimalarial drug artesunate (ARS) and its metabolite dihydroartemisinin (DHA) in human plasma was reported in 2008 [9]. It was based on precipitation of plasma proteins with acetonitrile, containing SIL forms of the analytes as ISs, and direct injection of the supernatant into an LC–MS/MS system. When the method was applied for the analysis of plasma from malaria patients, an extreme variability of the IS response was observed, as illustrated in Figure 3A for the IS for DHA. Since the IS response in calibrators and QCs was not affected, it was concluded that a matrix-related issue had to be the root cause of the observed variability, and that the composition of malaria patient plasma for some reason was responsible for the effect. This reasoning was supported by the fact that method validation, using plasma from healthy subjects, had been completed without any major variation in IS response, and by the observation that the analysis of plasma from patients who had recovered from malaria was also unaffected. The IS responses were only seen to decrease, so the authors concluded that the analytes and their ISs might very well degrade in the presence of disease-related compounds in plasma, a conclusion that is also obtained by following the decision tree in Figure 1.

It was already known that artemisinin analogs such as ARS and DHA are unstable in the presence of iron(II). In fact, the pharmacological mechanism of action of these compounds is based on this behavior. The malaria parasite *Plasmodium falciparum* consumes hemoglobin in red blood cells and releases iron(II)-heme complex into the blood stream, which reacts with artemisinin derivatives to a compound with antimalarial activity. Plasma samples from patients with different severities of malaria therefore contain varying levels of iron(II), which will result in a corresponding variability in the degree of degradation of ARS, DHA and their SIL ISs. Further investigation showed that the pharmacokinetic profiles for both ARS and DHA did not seem to be affected by the reduction in IS response. This means that the ratio of the analyte and IS concentrations apparently remained constant, which

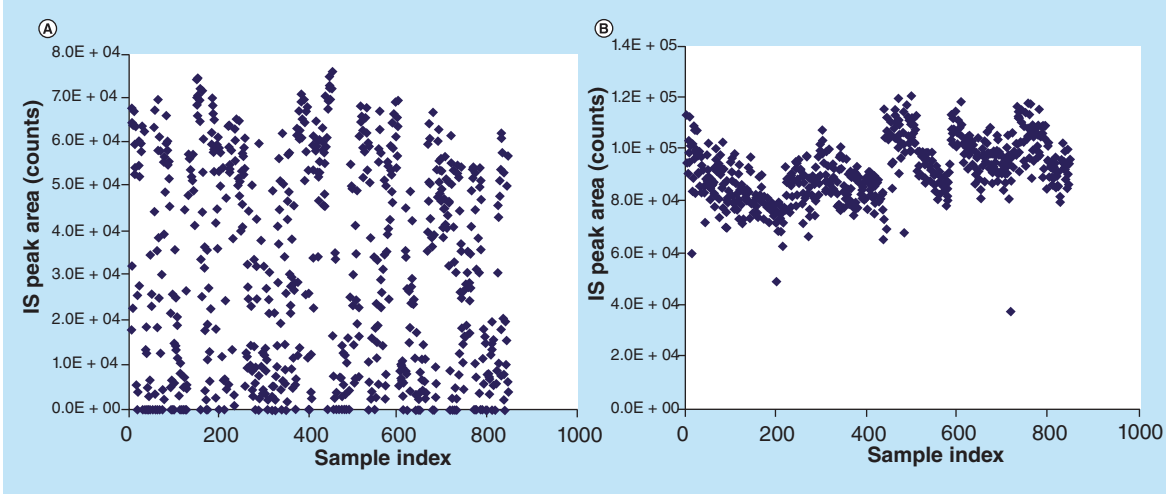

Figure 3. Response of the internal standard for dihydroartemisinin in the same set of malaria patient plasma samples. With protein precipitation **(A)** and solvent-free solid-phase extraction **(B)** as sample extraction approach.
IS: Internal standards.
Reproduced with permission from [9], © 2008 Elsevier B.V. (2008).

can only be the case if the degradation took place after addition of the ISs. Pre-analytical decomposition of the analytes during sampling or storage could therefore be excluded.

Since the analytes were stable in untreated patient plasma, and their instability appeared to be initiated by the protein precipitation step, it was speculated that the release of the analytes from plasma binding proteins by the addition of an organic solvent was responsible for exposing them to iron(II) and inducing their degradation. Indeed, replacing the protein precipitation step by a solvent-free solid-phase extraction largely reduced the variability of the IS response in the same set of patient plasma samples (Figure 3B). To keep the extraction completely solvent-free, the IS working solution was even prepared in a mixture of water and plasma instead of in an organic solvent. In this way, samples in which up to 2% of the original red blood cells had lysed could still be analyzed without a significant decrease of the response of either analyte or IS. For plasma with a higher degree of hemolysis, the IS response did decrease but the analyte/IS ratio remained unaffected.

In this case study, the root cause for the observed high IS variability was convincingly demonstrated. It was also concluded that, since analytes and ISs are affected to the same extent, the plasma concentration results can still be considered reliable as long as both the analyte and IS peak area do not drop below the normal value at the LLOQ level. In the original analysis, about 15% of the results were unquantifiable because of complete loss of response, so application of the improved method was needed to report all concentrations.

Case study 3: matrix-induced ionization enhancement

Olsson *et al.* reported highly variable IS responses in the LC–MS/MS analysis of mouse plasma samples for a proprietary compound [10]. The samples had been acidified at the time of collection and were later subjected to a simple protein precipitation extraction with acetonitrile in the presence of a ^{13}C-labeled IS. A portion of the supernatant was then evaporated to dryness followed by reconstitution and analysis by LC–MS/MS. Method validation had taken place without problems, but out of the 158 study samples that were analyzed, 22 exhibited a more than twofold increase in IS response, while calibrators and QCs spiked in control matrix were unaffected (Figure 4). As also suggested by Figure 1, an effect related to the composition of the matrix is likely to be responsible, but degradation as described in the previous case could be excluded, since the IS responses only increased. Through a series of experiments, the authors determined the likely root cause of the variable IS to be the endogenous bile acid taurocholic acid (TCA), that partially co-eluted with the analyte of interest and the IS, and most probably caused enhancement of the ionization efficiency of the IS. This conclusion was confirmed by the analysis of test solutions containing a constant amount of the IS and increasing concentrations of TCA, where it was observed that the IS response increased for higher TCA concentrations. In addition, the test solutions were also analyzed using chromatographic settings with different degrees of resolution between the IS and TCA. In the case of complete

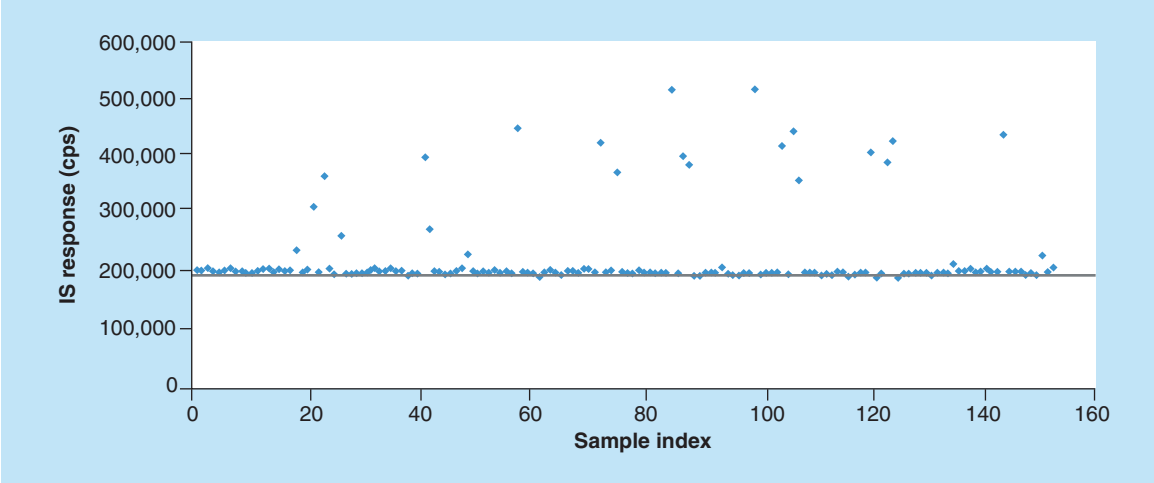

Figure 4. Response of the internal standard of a proprietary compound in mouse plasma.
IS: Internal standard.
Reproduced with permission from [10].

chromatographic resolution between the two, no significant IS response enhancement was observed, which supports the assumed ionization enhancement by TCA.

To investigate whether the accuracy of the results was affected, multiple individual samples were analyzed using two different injection volumes (i.e., 2 µl and the original 10 µl). Two of the samples that had not initially displayed an IS response enhancement were injected at the two volumes and showed the expected fivefold decrease of the peak areas of analyte and IS with no significant change in the area ratio. Conversely, for two samples that had originally shown an increased IS response, a disproportionate reduction in IS peak area was observed. For the lower injection volume, the IS response was now only marginally higher than for the other, previously unaffected samples. Since the peak area ratios remained unchanged also for these samples, analyte and SIL-IS were shown to behave similarly in case TCA co-elutes and thus quantification is not expected to be impacted. Therefore, all results were accepted and no samples were re-analyzed.

Interestingly, the authors also demonstrated the effect of the presence of elevated levels of TCA on the ionization behavior of some other drugs. Ionization efficiency was found to be modified in all cases, from 50% suppression to over 100% enhancement, which illustrates that the effect is strongly compound-dependent. While TCA levels are normally relatively low, it is obvious that physiological levels occasionally occur, which impact the ionization of analytes and ISs in case there is insufficient chromatographic separation. The example described in this paper shows the importance of testing several lots of the biological matrix when developing and validating a method, although it can never be ensured that such an effect will always be picked up. This case study also emphasizes the great importance of using proper ISs. Here, analyte and IS co-eluted were affected in exactly the same way. If retention times would have been only slightly different, such as in the case of an analog or even deuterium-labeled IS, no more full correction would have been provided and the reliability of the results would have been seriously compromised.

Case 4: storage-induced sample inhomogeneity

In our own laboratory, an LC–MS/MS method for a proprietary small molecule was transferred in from an external lab, where the method had been developed and validated without reported issues. Briefly, a SIL IS was added to an aliquot of plasma and, after pH adjustment, liquid extraction with ethyl acetate was performed in a 96-well format. Part of the organic solvent was transferred to a new 96-well plate, evaporated to dryness and reconstituted in a mixture of acetonitrile and water. The extract was analyzed with reversed-phase LC and positive-mode electrospray MS/MS. Method transfer and validation occurred without problems and typically with a low variation in IS response (see Figure 5A). However, upon reinjection of an extracted validation batch after storage of 120 h in the autosampler, a dramatic increase in IS response variability was seen (Figure 5B).

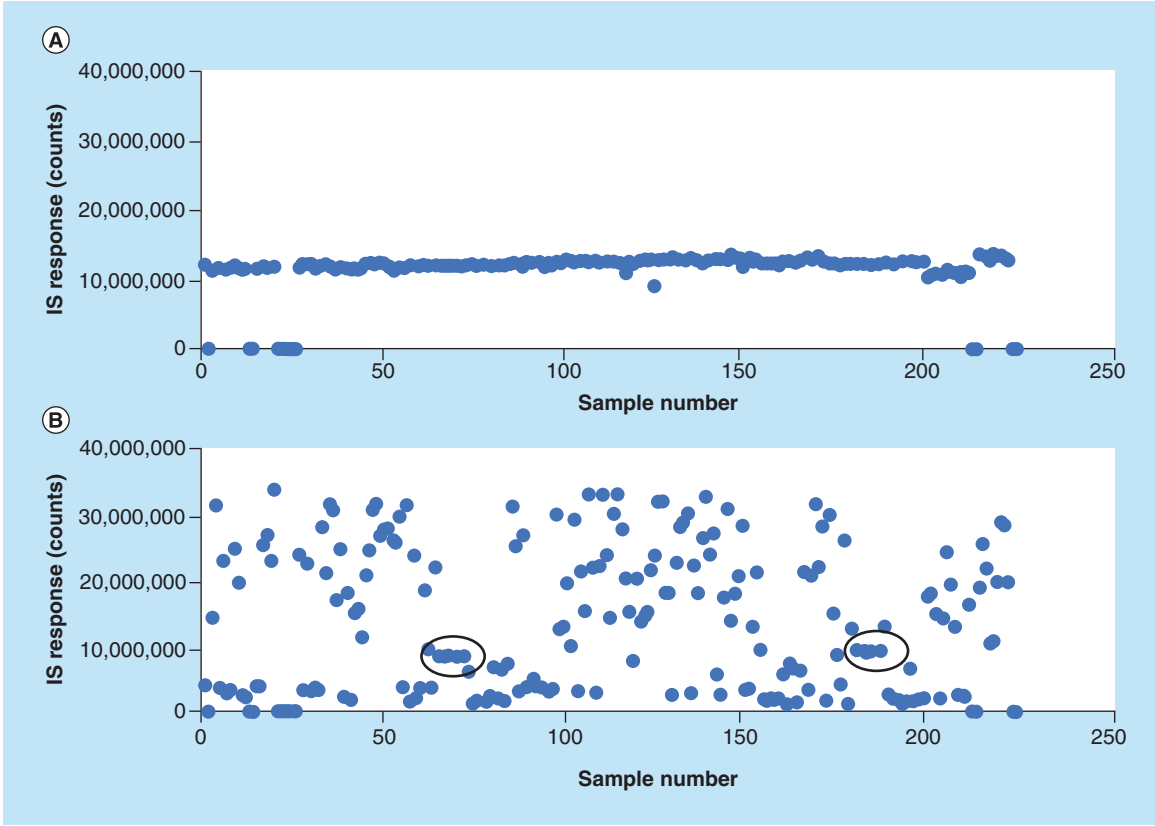

Figure 5. Response of the internal standard of a proprietary compound in the same set of plasma extracts. Original results, CV 5.3% **(A)** and re-injection results after storage in the autosampler, CV 78%. **(B)** Responses for test injections are indicated with circles.
IS: Internal standard.

Investigation into the root cause, by following the decision tree in Figure 1, led to a number of observations and conclusions. First, the reinjected batch only contained spiked samples, the majority of which had been prepared with a single lot of blank plasma. Differences in composition of the plasma matrix could, therefore, be ruled out as a cause for the effect. Since a normal IS response plot was obtained when samples were injected directly after extraction, the extraction recovery must originally has been consistent, and this suggests that the IS responses had changed as a result of storage in the autosampler. Clearly, the IS responses had both increased and decreased across the reinjected run, so degradation or solvent evaporation during storage cannot be held responsible for the observed variability, as these would only have led to a decrease or increase, respectively. For this reason, it was suspected that sample inhomogeneity could be responsible for the observed variability.

Despite the high and complex IS response variability, all calibration and validation results passed, which suggests that the IS was affected to the same extent as the analyte, which was not unexpected for a SIL form. Interestingly, a set of test solutions (matrix-free and not subjected to extraction) was part of the reinjected batch and their IS responses were essentially the same as for the original analyses. In addition, a small, but significant difference in retention time between plasma extracts and test solutions was seen in the reinjected run, which did not occur in the original run. This suggests that the composition of the plasma extract (but not the test solution) may have changed after storage. Since the amount of analyte and IS that was detected varied considerably and randomly from one sample to the other, it could be speculated that the extracts had become increasingly inhomogeneous upon storage, and that analyte and IS had accumulated in some parts of an extract sample, while other parts had become depleted.

The exact root cause remains unclear. Perhaps there happened to be remaining traces of ethyl acetate after solvent evaporation, which had dissolved in the highly aqueous reconstitution solvent at ambient temperature but started to separate upon storage at the lower temperature (10°C) in the autosampler. If analyte and IS accumulated in this ethyl acetate compartment, a lower or higher amount may have been injected depending on the exact composition

of the extract and the position of the injection needle. Alternatively, extracted matrix components may have started to precipitate upon storage, with the analyte (partly) incapsulated in the precipitate. Since method validation had already been completed and all results passed, no further modifications to the analytical methodology were done. To avoid the effect from happening, it was ensured during subsequent application of the method, that all samples were injected immediately after extraction.

Conclusion

In this paper, four different case studies were examined that exhibited extremely variable IS responses. For each case, a systematic approach helped finding the (probable) root cause and assessing the impact on the reliability of the results. As discussed in some detail, many factors can contribute to variable IS response, originating from suboptimal performance of the method and/or issues related to sample composition. Fortunately, analytical results are not typically affected by this variability since more often than not the IS behaves the same as the analyte and thus normalizes the results. The case studies also illustrate the importance of having an IS with a close similarity to the analyte, for LC–MS/MS preferably a SIL one. Still, it cannot be assumed that an IS will always compensate for fluctuations and too high a variability is usually a sign of poor method performance, which may be need to be improved before continuing with sample analysis. Similarly, just because the IS responses fall within a predefined acceptable range, it does not necessarily mean that one should not closely look at the data for trends, that may indicate analytical or sample-related issues and, thus, potentially incorrect results.

Future perspective

ISs have always been and will continue to be essential for high-quality bioanalytical LC–MS/MS methods. It is our expectation that a careful evaluation of IS responses, root cause investigation of clearly deviating values and the corresponding assessment of its impact on the accuracy of analytical results will increasingly become a routine procedure in bioanalytical laboratories.

Executive summary

- Highly variable internal standard responses can be due to a variety of (analytical and pre-analytical) causes.
- Much information about the root cause and the impact on data reliability often can be derived from existing results, by a logical step-by-step evaluation.
- Additional experiments may be helpful to confirm or further investigate the root cause and to assess the impact on data reliability.
- Highly variable internal standard responses are indicative of nonoptimal analytical and/or pre-analytical conditions and may call for method adjustment, even if data are concluded not to be affected.

Financial & competing interests disclosure

The authors have no relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript. This includes employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties.

No writing assistance was utilized in the production of this manuscript.

References

1. Wieling J. LC–MS/MS experiences with internal standards. *Chromatographia* 55, S107–S113 (2002).
2. Tan A, Boudreau N, Lévesque A. Internal standards for quantitative LC–MS bioanalysis. In: *LC–MS in drug bioanalysis*. Xu Q, Madden T (Eds). Springer, MA, USA (2012).
3. Stokvis E, Rosing H, Beijnen JH. Stable isotopically labeled internal standards in quantitative bioanalysis using liquid chromatography/mass spectrometry: necessity or not? *Rapid Commun. Mass Spectrom.* 19(3), 401–407 (2005).
4. Berg T, Strand DH. ^{13}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* 1218(52), 9366–9374 (2011).
5. Lowes S, Jersey J, Shoup R *et al.* Recommendations on: internal standard criteria, stability, incurred sample reanalysis and recent 483s by the Global CRO Council for Bioanalysis. *Bioanalysis* 3(12), 1323–1332 (2011).
6. White S, Adcock N, Elbast W *et al.* European Bioanalysis Forum: recommendation for dealing with internal standard variability. *Bioanalysis* 6(20), 2767–2774 (2014).

7. Welink J, Fluhler E, Hughes N *et al.* 2015 White Paper on recent issues in bioanalysis: focus on new technologies and biomarkers (part 1 – small molecules by LCMS). *Bioanalysis* 7(22), 2913–2925 (2015).
8. Tan A, Hussain S, Musuku A, Massé R. Internal standard response variations during incurred sample analysis by LC–MS/MS: case by case trouble-shooting. *J. Chromatogr. B* 877(27), 3201–3209 (2009).
9. Lindegardh N, Hanpithakpong W, Kamanikom B *et al.* Major pitfalls in the measurement of artemisinin derivatives in plasma in clinical studies. *J. Chromatogr. B* 876(1), 54–60 (2008).
10. Olsson AO, Jelvestam M, Ahnoff M. Variability of the IS revealed ionization enhancement by bile acid in mouse plasma. *Bioanalysis* 5(19), 2371–2378 (2013).