Analytica Chimica Acta xxx (2018) 1–11

Contents lists available at ScienceDirect

Analytica Chimica Acta

journal homepage: www.elsevier.com/locate/aca

Evaluation of batch effect elimination using quality control replicates in LC-MS metabolite profiling

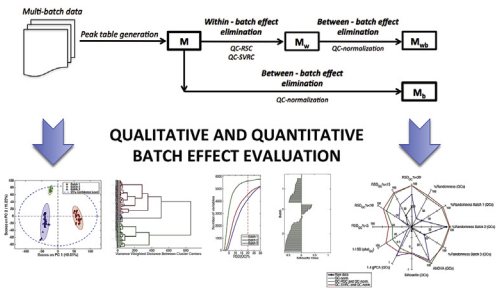
Ángel Sánchez-Illana [a], Jose David Piñeiro-Ramos [a], Juan Daniel Sanjuan-Herráez [b], Máximo Vento [c], Guillermo Quintás [b, d, *], Julia Kuligowski [a]

[a] Neonatal Research Unit, Health Research Institute La Fe, Valencia, Spain
[b] Health and Biomedicine, Leitat Technological Center, Barcelona, Spain
[c] Division of Neonatology, University & Polytechnic Hospital La Fe, Valencia, Spain
[d] Analytical Unit, Health Research Institute La Fe, Valencia, Spain

HIGHLIGHTS

- Quantitative/qualitative tools for the analysis of LCMS batch effects are reviewed.
- Graphical integration of outputs from multiple tools facilitates method optimization.
- Batch effect elimination improves data quality irrespectove of the approach.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 6 September 2017
Received in revised form
6 January 2018
Accepted 21 February 2018
Available online xxx

Keywords:
Metabolomics
Batch effect
QC-SVRC
QC-RSC
LC-MS

ABSTRACT

Systematic variation of the instrument's response both within- and between-batches is frequently observed in untarget LC-MS metabolomics involving the analysis of a large number of samples. The so-called batch effect decreases the statistical power and has a negative impact on repeatability and reproducibility of the results. As there is no standard way of assessing or correcting LC-MS batch effects and there is no single method providing optimal results in all situations, the selection of the optimal approach is not trivial.

This work explores the effectiveness of a set of tools for batch effect assessment. Qualitative tools include the monitoring of spiked internal standards, principal component analysis and hierarchical cluster analysis. Quantitative tools comprise the distribution of RSD_{QC} values, the median Pearson correlation coefficient in QCs, the ratio of *random* features in QCs using the runs test, as well as multivariate tools such as the δ-statistic, Silhouette plots, Principal Variance Component Analysis and the expected technical variation in the prediction. Results show that qualitative and quantitative approaches are complementary and that by limiting the analysis to QCs the power to detect and evaluate both within and between batch effects is increased. Besides, the graphical integration of outputs from multiple quantitative tools facilitates the evaluation of batch effects and it is proposed as a straightforward way for comparing and tailoring batch effect elimination approaches.

* Corresponding author. Health and biomedicine, Leitat Technological Center, Barcelona, Spain.
E-mail address: gquintas@leitat.org (G. Quintás).

https://doi.org/10.1016/j.aca.2018.02.053
0003-2670/

1. Introduction

Global metabolomic profiling aims at the measurement of the totality of small molecular weight endogenous metabolites present in a biological sample [1]. Due to the diversity in physicochemical properties and concentration among the metabolites, there is no single analytical technique enabling its complete profiling. Nonetheless, ultra performance high resolution liquid chromatography – mass spectrometry (LC-MS) is vastly employed in the field of untargeted metabolomic profiling [2] due to its superior sensitivity, selectivity and versatile detection capabilities.

Large scale clinical metabolomic studies typically show a high biological variability among subjects, which is especially relevant in case of the analysis of urine samples, where unspecific variations of the sample concentration of up to a factor of 10 due to changes in water uptake, diet, drugs or environmental factors are frequently observed [3], thus requiring carefully controlled experimental conditions and sample normalization [4]. LC-MS metabolomic data also show technical variation arising from different sources such as sample preparation or injection volume. Besides, LC-MS has specific sources of variation caused by e.g. inlet interface contamination or changes in the column performance [5,6] that introduce a gradual drift in the instrumental response during the analysis of a batch of samples (i.e. within-batch effect) including changes in e.g. retention time, peak shape, sensitivity or MS accuracy. In addition, systematic variation in the instrumental response when the analysis is split into several analysis batches (i.e. between-batch effect) may also occur due to e.g. system cleaning, calibration, differences in the mobile phase composition or MS detection sensitivity and resolution [7]. Instrumental variation hinders data interpretation, decreases the statistical power to identify metabolic responses and limits the repeatability and reproducibility of the results.

A number of approaches have been proposed and evaluated for both, between and within batch effect correction in the last years. The ComBat algorithm [8] has been employed in metabolomics for between-batch effect elimination. It is an empirical Bayes method that estimates the mean and variance for each metabolic feature by pooling information from multiple features with similar distributions in each batch [9]. However in unbalanced data sets, class differences may induce apparent batch differences and this type of batch effect correction should be used with caution [10]. Another recent cluster-based approach allows compensating multiple drift patterns within a batch. It relies on splitting the measurement batch into subsets which are then corrected separately by clustering of features, drift modelling and correction and removal of individual features with poor reproducibility between subsets [11]. A different, widely used approach for batch effect elimination is based on the repeated analysis of quality control (QC) replicates assuming that QC responses are independent of the injection order [12–14]. QCs are typically prepared by pooling samples to match the matrix composition and concentration ranges of the metabolites in the study population. Among others, the "Removal of Unwanted Variation" (RUV) approach [15] has been recently applied to batch effect normalization in metabolomics [13]. RUV is based on a multivariate model of the subspace of the unwanted variation in QCs obtained by principal component analysis (PCA). Then, the batch effect (i.e. unwanted variation) is estimated and subtracted from the original data of the study samples in this subspace. Alternatively, the change in the intensity of each metabolic feature (e.g. peak area) in QC replicates in dependence of the injection order can also be used to fit a function which can then be employed for data normalization of the study samples to the estimated trend [14,16–18]. This approach allows an effective handling of batch effects without modifying group effects, even in unbalanced experiments where the study groups are not evenly distributed across batches, providing at the same time a straightforward evaluation of both chromatographic (e.g. retention time, peak width) and MS (e.g. mass accuracy, detector response) performances. Two robust methods have been proposed for the fitting of the functions that mainly differ in the type of regression: Quality Control – Robust Spline Correction (QC-RSC) [16] and Quality Control – Support Vector Regression Correction (QC-SVRC) [18]. While QC-RSC uses an adaptive cubic splines function, QC-SVRC is a non-parametric approach that uses a radial basis function (RBF) kernel for support vector regression (SVR).

Besides batch effect elimination, data analysis workflows often include data cleaning steps for the elimination of uninformative features to reduce the likelihood of observing chance correlations and improve the precision, accuracy and interpretability of metabolic models [19,20]. Different methods commonly used include the Wilcoxon signed rank test to assess whether the QC distribution can be considered as representative of the total biological distribution [21], the selection of a threshold value (e.g. 20%) for the relative standard deviation in QCs (RSD_{QC}) [16] [22], or a minimum frequency of detection of a metabolic feature (e.g. 80% of the samples) [23]. A very effective method suitable when sample replicates are available is the comparison of between and within individual variances per metabolic feature for the assessment of the batch effect and the identification of peaks providing no biological information [16].

In summary, after peak table generation LC-MS metabolomic data still requires extensive pre-treatment to account for non-biological sources of variation. However, there is no standard way of assessing or correcting batch effects in LC-MS data and the selection of the optimal approach is not trivial as there is no single method providing optimal results in all situations. This work describes a set of tools for the evaluation of batch effects divided into qualitative and quantitative utilities [9]. Qualitative tools include the monitoring of spiked internal standards, PCA and hierarchical cluster analysis (HCA). Quantitative tools comprise the distribution of RSD_{QC} values, the median Pearson correlation coefficient in QCs, the ratio of *random* features in QCs using the *runs* test, as well as multivariate tools such as guided PCA [24,25], Silhouette plots, Principal Variance Component Analysis (PVCA) [26,27] and the technical variation in the prediction. These univariate and multivariate tools provides a way to detect low quality batches as well as a basis for the selection and improvement of batch effect elimination approaches. To illustrate each tool, data acquired during the UPLC-ESI(+)-TOF-MS analysis of 316 urine and 76 QC samples in three batches was employed.

2. Materials and methods

2.1. Sample collection and preparation

Urine samples were collected from newborns in the frame of a multi-center randomized clinical trial previously approved by the Ethics Committee of the Hospital Universitario y Politécnico La Fe (Valencia, Spain) (1645-CI-058). 700 μl of urine were collected from newborns within 96 h after birth and stored at −80 °C until analysis. Samples were thawed on ice and centrifuged. 50 μl of supernatant were withdrawn and spiked with 50 μl of IS solution. Two different IS solutions were used (i.e. IS_0 and IS_1) and the IS solution employed was randomly selected for each patient to create a controlled class difference between two groups of samples. Both IS spiking solutions included phenylalanine-D_5 ($PheD_5$, Cambridge Isotopes Laboratory Inc., Andover, MA, USA), caffeine-D_9 ($CaffD_9$, Toronto Research Chemicals, Toronto, Ontario, Canada), leukine enkephalin (LeuEnk) (Sigma-Aldrich Química SA, Madrid, Spain) and reserpine (Reserp) (Sigma-Aldrich Química SA) in H_2O:CH_3OH

Please cite this article in press as: Á. Sánchez-Illana, et al., Evaluation of batch effect elimination using quality control replicates in LC-MS metabolite profiling, Analytica Chimica Acta (2018), https://doi.org/10.1016/j.aca.2018.02.053