CHAPTER IV
RESULTS AND DISCUSSION

4.1 Development and Validation of Analytical Method

4.1.1 The Optimization of MS Conditions

The electrospray ionization source (ESI) interface was operated in the negative ion mode to generate, focus and transport the target anions of all eight acids to the ion-trap mass analyzer. This technique offered atmospheric pressure ionization at a low temperature and therefore an efficient ionisation of polar compounds. The ESI conditions were optimized (see section 3.1.4.1) for an efficient production of a fine aerosol. The ion trap MS parameters (also see section 3.1.4.1) were optimized to generate and accumulate the ions to achieve maximum responses of the target ions with the best mass accuracy and resolution. All MS conditions were optimized for each individual compound and implemented in the combined method for all eight acid standards. The optimized MS condition is shown in Table 2.

Table 2 Optimized ESI and MS parameters for eight acids.

ESI parameters	Capillary voltage	3000 V
	Nebulizer pressure	40 psi
	Drying gas flow	8 L/min
	Drying gas temperature	365 ^{o}C
MS parameters	Capillary exit offset	-16.4 V
	Skimmer 1	-15.6 V
	Skimmer 2	-5.2 V
	Octapole	-1.5 V
	Len 1	1.0 V
	Len 2	23.0 V
	Trap drive	22.8

4.1.2 The Optimization of LC Conditions

All eight organic acids were soluble in water with pKa values ranging from approximately 2.22 to 4.41 [42]. Earlier methods for small polar compounds were GC-MS. However, GC-techniques require samples to be volatile and have thermal stability, and some samples must be derivatized prior to analysis [9, 43, 44]. Furthermore, derivatizing processes are often use toxic chemicals and time consuming. Ion-exchange chromatography requires mobile phases that usually contain high levels of aqueous solvents which could reduce the efficiency of the electrospray ionization (ESI) and thus the analytical sensitivity [10, 45]. Because of the polar properties of target analytes, a HILIC method was evaluated. HILIC can efficiently retain and separate small polar compounds, which are not retained in conventional reversed-phase LC [46]. HILIC also requires a high organic content in the mobile phase which potentially increases the desolvation efficiency and thus enhances the MS sensitivity. For LC optimization, the separation efficiency (also express as peak shape), the highest sensitivity (express as peak intensity) and the reproducibility were considered.

4.1.2.1 HILIC Condition in The Study

The LC separation was performed on a ZIC®-HILIC (250 mm × 2.1 mm, 5 µm) protected by a ZIC®-HILIC guard column (20 mm × 2.1 mm, 5 µm). The stationary phase of ZIC®-HILIC column is a porous silica particle covalently boned with highly polar sulfobetaine type zwitterionic functional group. The column temperature was controlled at 30°C and the injection volume was 5 µL. A mobile phase program for separation and washing column started with mobile phase A at flow rate 0.5 mL/min under isocratic conditions for 13 minutes. The column was then washed with mobile phase B at a flow rate 0.4 mL/min for 7 minutes. Finally, the column was re-equilibrated for 1 minute with the starting conditions before each new injection (total run time 21 minutes). Parameters used in this method are described hereafter.

4.1.2.2 Type, Concentration and pH of Mobile Phase

In HILIC mode, semi-aqueous mobile phases are employed. The mobile phase should be volatile with high organic content and could dissolve very polar analytes. Typically, mobile phase for HILIC consist of 40-97% MeCN in water or volatile buffer. The recommended ionic strength for the buffer in mobile phases is commonly between 5 to 20 mM but a 100 mM buffer provided a significant improvement in peak shapes (data not shown). Furthermore, a suitable pH range for silica-based HILIC columns is usually pH 3-8. According to a good miscibility with water, providing good HILIC retention and suitable viscosity, the mixture of MeCN and ammonium acetate 100 mM at pH 4.7 buffer, (80:20, v/v) was used as a mobile phase A for separation in this method.

4.1.2.3 The Wash Step Condition

Considering calibrators and QC, there were 73 biological samples (plasma/urine) to be analyzed in one batch and runtime for each sample was 21 minutes. The wash step was crucial to avoid endogenous interferences from plasma and urine. Furthermore, the wash step improved the reproducibility of the method and enabled the possibility of a high throughput analysis without degradation of the chromatographic performances or rise in the column back-pressure over an extended

period of time. Acetonitrile: Ammonium acetate 25 mM at pH 4.7, (50:50, v/v) was used for wash step with a higher aqueous ratio than the starting mobile phase but also a lower buffer ionic strength (less elution conditions).

4.1.2.4 Flow Rate and Runtime

The eight acids were separated under isocratic condition. Usually, compounds with longer retention times present a peak broadening in isocratic conditions. Therefore, flow rate and runtime should be considered and optimized to improve separation efficient (peak shape) of all acids. According to Figure 14, aKGA was the strongest retain acid in ZIC®-HILIC column and showed broad peak at flow rate 0.3 mL/min of mobile phase A. However, after increase flow rate of mobile phase A to 0.5 mL/min, aKGA had better peak shape and runtime of each sample was decreased (see Figure 15).

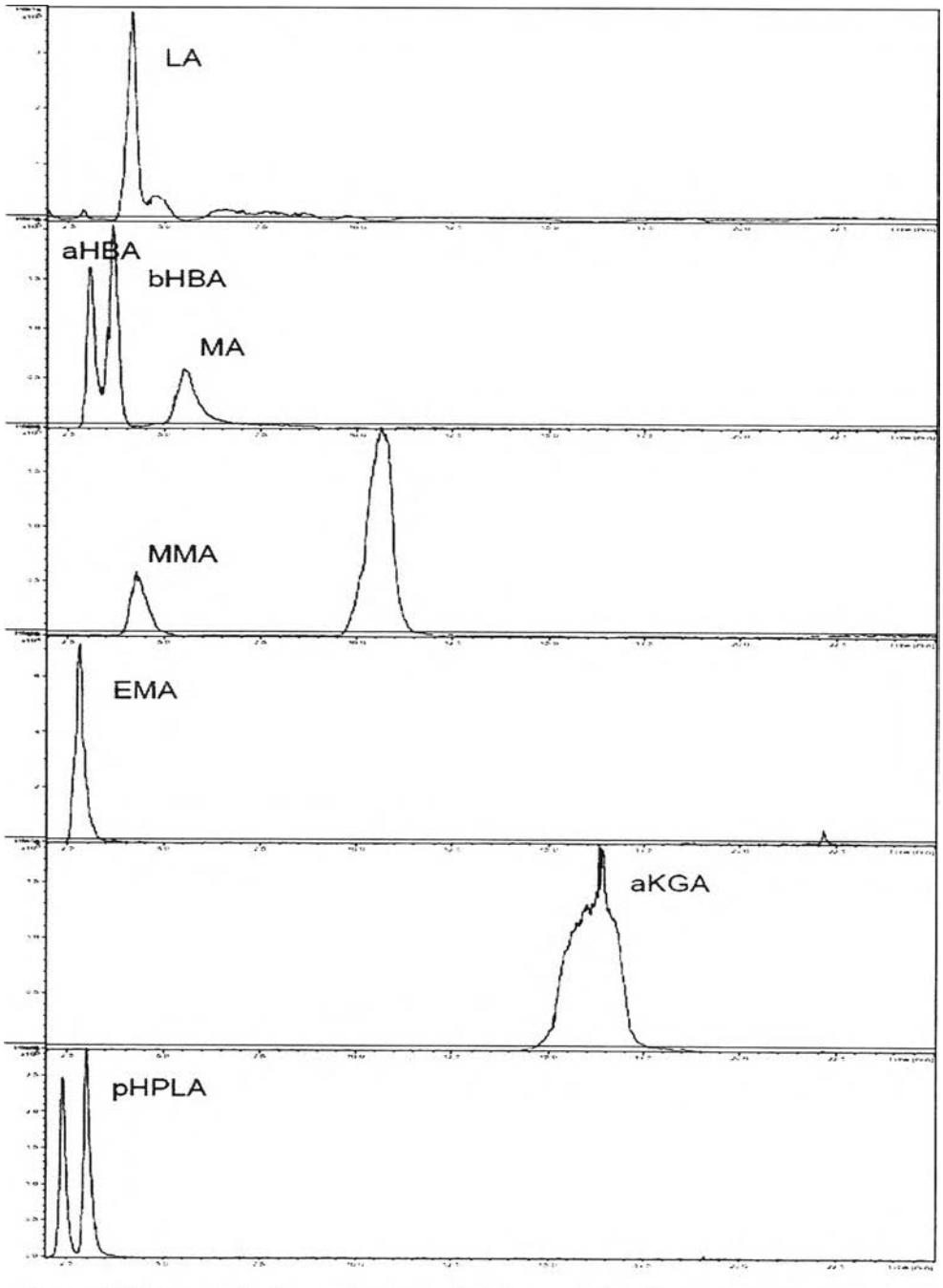

Figure 14 Extracted chromatogram of eight acids (at flow rate 0.3 mL/min of mobile phase A under HILIC condition in this study)

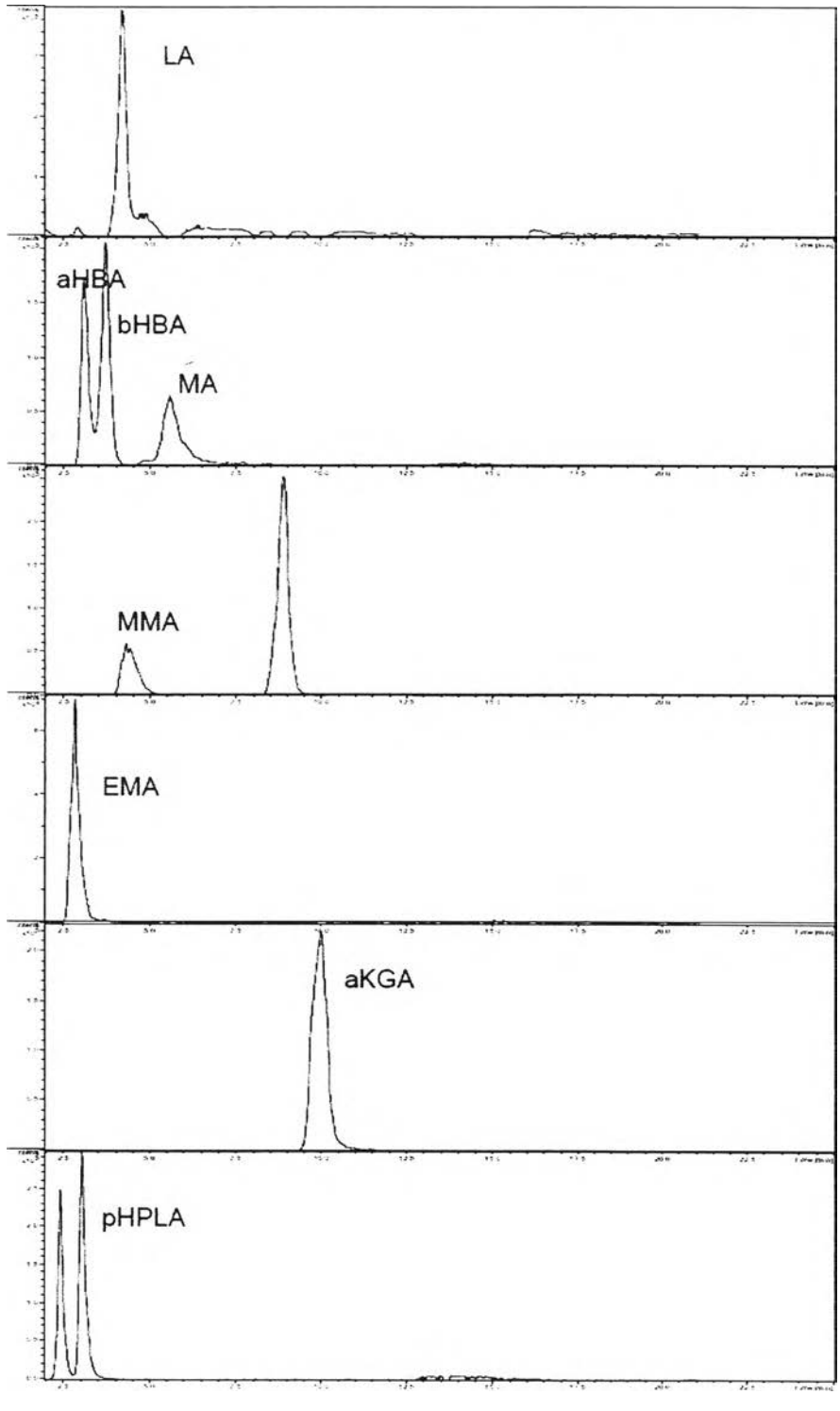

Figure 15 Extracted chromatogram of eight acids (at flow rate 0.5 mL/min of mobile phase A under HILIC condition in this study)

4.1.3 The Optimization of Extraction Method

The ion-exchange SPE offered cleaner extracts which led to minimising the effect of the matrix on the signal of the compounds investigated [47]. Smaller volumes are also used with SPE and the technique can be automated using a liquid handler for routine analysis in the future. In this study, ISOLUTE PE-AX 96-well SPE plate was used to extract anions (from small organic acids) in strong anion exchange mode from aqueous samples (see procedure in section 3.2.2). In addition, the high throughput extraction of 96 samples (including calibrators and quality controls) were achieved in one extraction batch.

The SPE optimization focused on optimizing the content of formic acid in the elution solvent to elute all small polar acids with the highest and most stable recoveries and also maintain appropriate evaporation time. Therefore, process efficiency (PE) and recovery (RE) (section 3.1.6.3) of all acids were compared between using 5, 10, 15 and 20 % formic acids in methanol as elution solvent. Furthermore, time of evaporation was also considered.

% PE and % RE of all seven acids (except aKGA) were in the range of 85-95% with %CV ≤10 when using 5, 10, 15 and 20 % formic acids in methanol as elution solvent. However, % PE and % RE of aKGA were increased in higher % formic acids content in methanol. They were increased from approximate 5.5 % to 95 % with %CV ≤8, when using 5 to 15 % formic acid in methanol. Furthermore, using 20 % formic acid in methanol did not increase %PE and %RE but extend time for evaporation. Therefore, the optimal elution solvent in this study was 15% formic acid in methanol (see the results for %PE and %RE in Table 5).

4.1.4 Method Validation for Determination of Endogenous Compounds in Biological Samples

4.1.4.1 Selectivity

The chromatograms of all eight acids in blank plasma and urine from healthy volunteers did not show any signals from interference in samples (i.e. concomitant medication), they showed only the endogenous analyte peaks (i.e. peaks that are found naturally in blank biological samples) within the retention time windows of the chromatographic peaks of analytes and IS. This was also verified

during the routine analysis. Average retention times (t_R) and extract target mass (m/z) of all analytes in six spiked healthy individuals plasma/urine are shown in Table 3.

4.1.4.2 Linearity of Calibration Curves and Sensitivity

A linear regression model with $1/x^2$ weighting resulted in the best accuracy (93.1-104.0%) and precision (<5.5%) over the entire calibration range and was considered the most appropriate regression model for all eight acids. This weighting expected considering the heteroscedasticity of data when validating a method over a broad calibration range. The final regression model also resulted in small and evenly distributed residual errors and high coefficient of regression. The accuracy of all eight acids ranged from 92.0-104.0%, 96.0-105.0% and 97.1-104.0% at LLOQ, ULOQ and over-curve, respectively. Inter-assay, intra-assay, and total precision of all eight acids at LLOQ, ULOQ and over-curve were <7.0%, <4.1% and <3.6%, respectively. There were no carry-over effects for any analytes or SIL-IS. The LOD, LLOQ and the calibration ranges of all eight acids are shown in Table 3.

The calibration curves were prepared in water and additional validation procedures were necessary to evaluate the appropriateness of this approach. At QC low, medium and high concentrations, the RF ranged from 0.93 to 1.04 (RSD<3.9%) in plasma and from 0.97 to 1.01 (RSD<1.4%) in urine for all eight acids. The implementation of RF for QC samples spiked in plasma/urine did not result in better accuracy or precision; i.e. the absolute sum of relative residuals (%RR) was higher when implementing a RF than without the correction for both plasma and urine samples. Absolute sum of %RR of all acids in plasma without RF was 46.3 while with RF was 79.9. In addition, %RR of all acids in urine without RF was 77.3 whereas with RF was 86.0. Thus, a RF was not necessary despite using different matrices for calibration curves and QC samples compared to clinical samples. Furthermore, adequate accuracy and precision were shown when spiking QC samples in blank plasma and urine and quantifying them using a calibration curve prepared in water (Table 4). This supports further the parsimonious approach of not compensating for the different matrices.

4.1.4.3 Recovery and Matrix Effects

All analytes showed high process efficiency and recovery with no substantial matrix effects in six individual healthy plasma/urine sources: process efficiency = 94.6-99.8% (RSD<2.7%); recovery = 94.0-99.8% (RSD<4.1%) and matrix effects = -3.7-3.4% (RSD<4.2%) (Table 5). The high recovery and lack of matrix effect confirmed that the SPE procedure provided an excellent method for extraction of the target analytes in both urine and plasma samples. Post-column infusions (Figure 16) showed that only lactic acid displayed a clear endogenous peak in the extracted blank plasma but not in urine, due to higher concentrations in blank plasma from healthy volunteers compared to the LOD of the assay.

4.1.4.4 Precision and Accuracy

The back-calculated concentrations of QC samples prepared in water resulted in acceptable accuracy of 96.4-100.1% with intra-assay, inter-assay and total precisions below 7.6% for all samples (Table 6).

4.1.4.5 Stability

All eight acids remained stable in plasma and urine during three freeze/thaw cycles, at ambient temperature, and also when stored at 4 °C up to 48 h. All acids were also stable at ambient temperature for at least 4 h before SPE, and at 20 °C in the autosampler for at least 72 h. All acids showed long-term stability in plasma and urine matrices for at least 1 month, and in stock solution (water) for at least 4 months. All stability results met FDA acceptance criteria, the differences between the stored and fresh solution of all eight acids in plasma and urine were <8.6% with a precision <7.0%.

Table 3 Method specification and calibration range of eight organic acids

Acids	Molecular weight	m/z	t_R	LOD	LLOQ	Method calibration range
	(Da)		(min)	(µg/mL)	(µg/mL)	(µg/mL)
LA	90.1	89.1	4.30	0.750	2.50	2.50-2500
aHBA	104.1	103.0	3.10	0.0375	0.125	0.125-125
bHBA	104.1	103.0	3.80	2.25	7.50	7.50-375
pHPLA	182.2	180.9	3.10	0.0300	0.100	0.100–100
MA	104.1	103.0	5.70	0.300	1.00	1.00–1000
MMA	118.1	117.0	4.60	0.0750	0.250	0.250–250
EMA	132.1	131.0	3.00	0.0750	0.250	0.250–100
aKGA	146.1	144.9	11.5	9.00	30.0	30.0-1500

LA: L-lactic acid; aHBA: α-hydroxybutyric acid; bHBA: β-hydroxybutyric acid; pHPLA: p-hydroxyphenyllactic acid; MA: malonic acid; MMA: methylmalonic acid; EMA: ethylmalonic acid; aKGA: α-ketoglutaric acid; t_R: retention time; m/z: target mass; LLOQ: lower limit of quantification; LOD: limit of detection

Table 4 Accuracy and precision of quality control samples spiked in human blank plasma and urine (n=5) analyzed at 3 separate occasions

Acids	QC level (µg/mL)	Accuracy (%)	Precision (%RSD)		
			Plasma		
			Inter-assay	Intra-assay	Total-assay
LA	QC1, 7.50	97.9	5.5	2.5	3.1
	QC2, 1252	100.4	6.1	0.6	2.4
	QC3, 2252	100.5	0.7	0.6	0.6
aHBA	QC1, 0.375	96.2	5.2	2.6	3.1
	QC2, 62.6	97.4	10.2	1.5	4.1
	QC3, 113	95.3	6.7	1.3	2.8
bHBA	QC1, 22.5	93.9	11.5	3.5	5.4
	QC2, 188	101.2	3.1	1.6	1.9
	QC3, 338	100.1	4.7	0.9	2.0
pHPLA	QC1, 0.30	95.9	5.9	2.7	3.3
	QC2, 50.0	99.1	4.3	2.4	2.7
	QC3, 90.0	101.6	2.9	0.9	1.4
MA	QC1, 3.00	97.1	6.4	1.6	2.8
	QC2, 501	99.7	4.2	0.5	1.6
	QC3, 901	99.8	1.3	0.4	0.6

MMA	QC1, 0.75	99.6	2.6	1.5	1.7
	QC2, 125	97.7	7.6	1.1	3.0
	QC3, 225	99.9	3.1	1.0	1.5
EMA	QC1, 0.75	95.4	4.1	2.3	2.6
	QC2, 50.1	103.4	4.1	1.2	1.9
	QC3, 90.1	100.6	0.9	1.0	1.0
aKGA	QC1, 90.0	96.8	6.1	1.8	2.8
	QC2, 751	100.1	0.2	0.4	0.4
	QC3, 1351	98.9	4.8	0.4	1.9

Table 4 (con) Accuracy and precision of quality control samples spiked in human blank plasma and urine (n=5) analyzed at 3 separate occasions

			Urine		
Acids	QC level (µg/mL)	Accuracy (%)	Precision (%RSD)		
			Inter-assay	Intra-assay	Total-assay
LA	QC1, 7.50	97.1	3.7	2.4	2.6
	QC2, 1252	100.6	2.1	0.5	0.9
	QC3, 2252	100.3	1.8	0.2	0.7
aHBA	QC1, 0.375	95.9	12.1	2.4	5.1
	QC2, 62.6	96.4	4.8	2.3	2.8
	QC3, 113	99.9	5.4	3.0	3.4

bHBA	QC1, 22.5	97.2	1.1	3.3	3.0
	QC2, 188	101.0	7.0	1.7	3.0
	QC3, 338	100.5	1.3	1.4	1.4
pHPLA	QC1, 0.30	98.6	7.5	1.9	3.3
	QC2, 50.0	98.1	5.2	2.1	2.8
	QC3, 90.0	101.4	3.9	1.7	2.1
MA	QC1, 3.00	95.9	7.5	1.5	3.1
	QC2, 501	100.4	0.5	0.5	0.5
	QC3, 901	100.0	0.9	0.3	0.5
MMA	QC1, 0.75	97.7	4.9	1.6	2.3
	QC2, 125	97.1	4.9	1.0	2.1
	QC3, 225	99.3	0.5	0.8	0.8
EMA	QC1, 0.75	98.1	7.3	1.7	3.2
	QC2, 50.1	101.8	5.2	1.0	2.2
	QC3, 90.1	100.9	2.2	0.6	1.0
aKGA	QC1, 90.0	93.5	5.2	1.8	2.6
	QC2, 751	100.0	0.6	0.8	0.8
	QC3, 1351	100.1	0.1	0.5	0.4

Table 5 Process efficiency, recovery and matrix effects in human plasma (n=6) and urine (n=6)

Acids	QC level (μg/mL)	Plasma PE (%)	Plasma RE (%)	Plasma ME (%)	Urine PE (%)	Urine RE (%)	Urine ME (%)
LA	QC1, 7.50	99.2 ± 1.7	97.3 ± 3.3	2.0 ± 3.52	96.0 ± 2.4	95.3 ± 2.4	0.8 ± 2.7
	QC2, 1252	97.9 ± 0.9	98.5 ± 0.6	-0.6 ± 1.5	95.6 ± 2.7	94.0 ± 1.9	1.7 ± 4.1
	QC3, 2252	95.8 ± 2.1	98.4 ± 1.0	-2.6 ± 2.8	97.5 ± 1.4	98.9 ± 1.1	-1.4 ± 1.1
aHBA	QC1, 0.375	98.7 ± 0.7	98.3 ± 0.8	0.4 ± 0.8	98.6 ± 1.1	99.5 ± 0.5	-0.9 ± 0.8
	QC2, 62.6	95.4 ± 2.3	97.6 ± 1.0	-2.2 ± 2.1	98.9 ± 1.1	98.6 ± 0.8	0.3 ± 0.9
	QC3, 113	96.5 ± 1.9	98.4 ± 0.8	-1.9 ± 1.5	98.8 ± 0.6	98.7 ± 1.1	0.1 ± 1.4
bHBA	QC1, 22.5	98.1 ± 0.9	98.6 ± 0.7	-0.5 ± 1.1	99.3 ± 1.2	98.3 ± 1.2	1.0 ± 2.1
	QC2, 188	97.1 ± 2.0	94.8 ± 1.1	2.4 ± 1.8	99.0 ± 0.9	97.1 ± 1.1	1.9 ± 1.5
	QC3, 338	95.9 ± 1.8	97.0 ± 2.0	-1.1 ± 3.1	98.1 ± 1.0	99.2 ± 0.9	-1.1 ± 1.1
pHPLA	QC1,0.30	95.1 ± 1.7	94.4 ± 2.6	0.7 ± 2.8	98.7 ± 2.0	98.0 ± 1.2	0.8 ± 1.4
	QC2, 50.0	94.6 ± 2.4	98.2 ± 2.1	-3.7 ± 2.0	98.6 ± 1.5	99.4 ± 0.6	-0.8 ± 1.3
	QC3, 90.0	97.0 ± 1.5	99.6 ± 1.2	-2.7 ± 0.7	99.2 ± 0.4	99.3 ± 0.4	-0.1 ± 0.7
MA	QC1, 3.00	97.9 ± 1.5	94.9 ± 0.9	3.2 ± 2.0	99.1 ± 0.4	99.5 ± 1.0	-0.4 ± 0.8
	QC2, 501	101.2 ± 0.7	99.0 ± 1.9	2.3 ± 1.5	98.9 ± 0.6	99.0 ± 1.1	-0.1 ± 1.4
	QC3, 901	98.8 ± 1.2	99.2 ± 1.3	-0.4 ± 1.6	99.5 ± 0.7	98.6 ± 1.1	1.0 ± 1.5
MMA	QC1, 0.75	99.8 ± 1.6	99.8 ± 2.3	0.1 ± 1.4	99.6 ± 0.8	98.9 ± 0.9	0.7 ± 0.9
	QC2, 125	96.9 ± 0.6	98.9 ± 0.6	-2.0 ± 1.2	98.7 ± 0.7	98.5 ± 0.8	0.1 ± 1.2
	QC3, 225	97.0 ± 2.4	98.8 ± 4.0	-1.7 ± 1.8	97.8 ± 1.7	98.6 ± 0.9	-0.8 ± 1.7

EMA	QC1, 0.75	97.2 ± 2.1	97.5 ± 2.6	-0.3 ± 2.5	97.3 ± 2.3	97.1 ± 2.5	0.2 ± 2.0
	QC2, 50.1	99.4 ± 0.7	98.7 ± 1.2	0.7 ± 0.8	98.9 ± 1.9	95.7 ± 1.1	3.4 ± 2.3
	QC3, 90.1	97.7 ± 1.0	95.0 ± 2.2	2.9 ± 3.0	99.0 ± 0.7	97.2 ± 1.2	1.8 ± 1.9
aKGA	QC1, 90.0	97.7 ± 2.1	97.4 ± 0.8	0.4 ± 1.9	96.7 ± 1.6	98.4 ± 1.1	-1.8 ± 1.4
	QC2, 751	96.1 ± 2.6	98.0 ± 1.7	-1.9 ± 2.6	95.7 ± 1.5	98.1 ± 1.3	-2.4 ± 2.5
	QC3, 1351	95.6 ± 2.7	99.1 ± 1.0	-3.5 ± 2.1	98.5 ± 1.2	98.2 ± 1.0	0.3 ± 1.6

Values are shown as average ± standard deviation; PE: process efficient; RE: recovery; ME: matrix effects

Table 6 Accuracy and precisions of QC samples in water for all eight acids (n=5) analyzed at 3 separate occasions

Acids	QC level (µg/mL)	Accuracy (%)	Precision (%RSD)		
			Inter-assay	Intra-assay	Total-assay
LA	QC1, 7.50	97.5	5.6	2.1	3.4
	QC2, 1252	97.6	5.7	2.8	3.8
	QC3, 2252	99.4	6.0	2.1	3.6
aHBA	QC1, 0.375	98.6	7.0	3.0	4.5
	QC2, 62.6	100.0	5.3	2.6	3.5
	QC3, 113	98.9	4.6	1.8	2.9
bHBA	QC1, 22.5	100.1	2.6	3.1	3.0
	QC2, 188	97.7	5.0	3.8	4.1
	QC3, 338	98.9	3.0	2.4	2.6
pHPLA	QC1, 0.3	97.1	4.4	1.9	2.8
	QC2, 50	98.7	3.6	2.0	2.5
	QC3, 90	98.3	4.4	1.5	2.6
MA	QC1, 3	98.3	2.3	3.5	3.2
	QC2, 501	97.4	4.1	1.9	2.7
	QC3, 901	98.5	2.9	1.7	2.1

MMA	QC1, 0.75	97.2	2.3	3.8	3.4
	QC2, 125	97.1	5.8	2.2	3.5
	QC3, 225	99.2	7.6	1.9	4.3
EMA	QC1, 0.75	98.2	3.5	2.0	2.5
	QC2, 50.1	99.8	2.6	2.3	2.4
	QC3, 90.1	100.7	4.3	1.7	2.7
aKGA	QC1, 90	97.0	3.7	2.1	2.6
	QC2, 751	96.4	2.3	0.9	1.4
	QC3, 1351	99.8	0.7	1.8	1.6

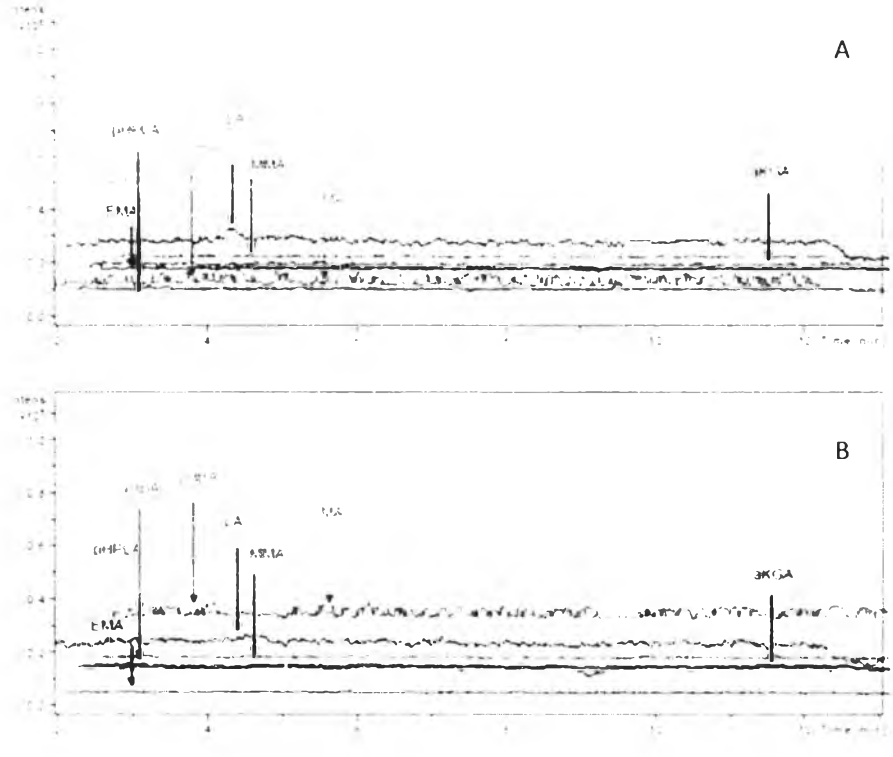

Figure 16 Extracted chromatogram of injected blank plasma (A) and urine (B) samples during post-column infusion of the 8 acid standards

(Arrows indicate expected retention time of analytes)

4.2 Clinical Applicability and Relevance

The clinical applicability of the analytical method was demonstrated by a preliminary analysis of plasma and urine samples collected in patients with severe malaria and uncomplicated malaria from Chittagong, Bangladesh and in healthy volunteers. The results showed that concentration of five of the eight potential acids were significantly elevated in plasma and seven were elevated in urine samples, respectively. An example of extracted chromatogram of a plasma and urine sample from a severe malaria patient recruited in the study is shown in Figure 18.

However, a complete study (section 3.3.1) showed that aKGA was found in only < 3% of all plasma samples from severe malaria patients, and not found in uncomplicated malaria and healthy plasma samples. Therefore, only four of the eight acids in plasma samples were considered and calculated in section 3.3 (statistical data analysis). The scatter plots (medians with interquartile range) of four acid and seven acid concentration found in plasma and urine samples from three sample groups shows that they are non-normal distribution for all acids in all groups (Figure 19 and Figure 20), respectively. A larger study has been planned to explore and evaluate in more detail the role of these acids in the pathogenesis of severe malaria.

The clinical relevance of five acids found in plasma presents that LA is the main contributor of acidosis in severe falciparum malaria [3, 6, 7] and this acid relates to Glycolysis pathway in Malaria parasite metabolic pathway [48]. Acidosis in severe malaria is caused by the accumulation of organic acids (including LA) and acids relate to ketoacidosis [3]. aHBA is the acid often found in patients suffering from lactic acidosis and ketoacidosis [49]. An increased ratio of NADH2/NAD in Pyruvate and LA equilibrium (Figure 17) is the most important factor for aHBA production [49, 50].

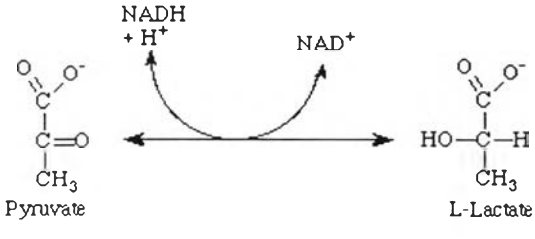

Figure 17 Equation of Pyruvate and L-Lactate (LA) equilibrium [50]

bHBA is a ketone body found in diabetic ketoacidosis [49]. This acid relates to Hypoglycaemia (an abnormally low blood glucose condition) which is associated with lactic acidosis. Hypoglycaemia is caused by a liver failure of Gluconeogenesis (metabolic pathway which causes the production of glucose from non-carbohydrate carbon substrates (i.e. pyruvate and lactate)) [3, 49]. Blood level of bHBA could be used for monitoring of diabetic patients [49, 51]. One of the most interesting acids in this study is pHPLA which seem to restricted to severe malaria. pHPLA is produced by the parasite and relate to Phenylalanine and Tyrosine metabolism in malaria parasite metabolic pathways [48, 49]. aKGA is the keto acid which relates to catabolic state and starving. This acid is a key intermediate in the Krebs cycle and be associated with Glutamate metabolism in Malaria parasite metabolic pathway [48, 49]. The level of aKGA in blood also relate to liver disease and coma [52].

Furthermore, clinical relevance of two other acids which found in only urine samples (not in plasma) shows that MMA is a derivative of MA (acid not found in both plasma and urine in this study). MMA is an important intermediate in metabolism of fat and protein. Abnormality of MMA metabolism causes Methylmalonic academia [49]. EMA is related to Ethylmalonic academia which is associated with an amino acid variation of short chain acyl-coenzyme A dehydrogenase [49]. Methylmalonic and Ethylmalonic academia are the metabolic disorders which disturb normal amino acid metabolism resulting in production of acids that usually not present [53].

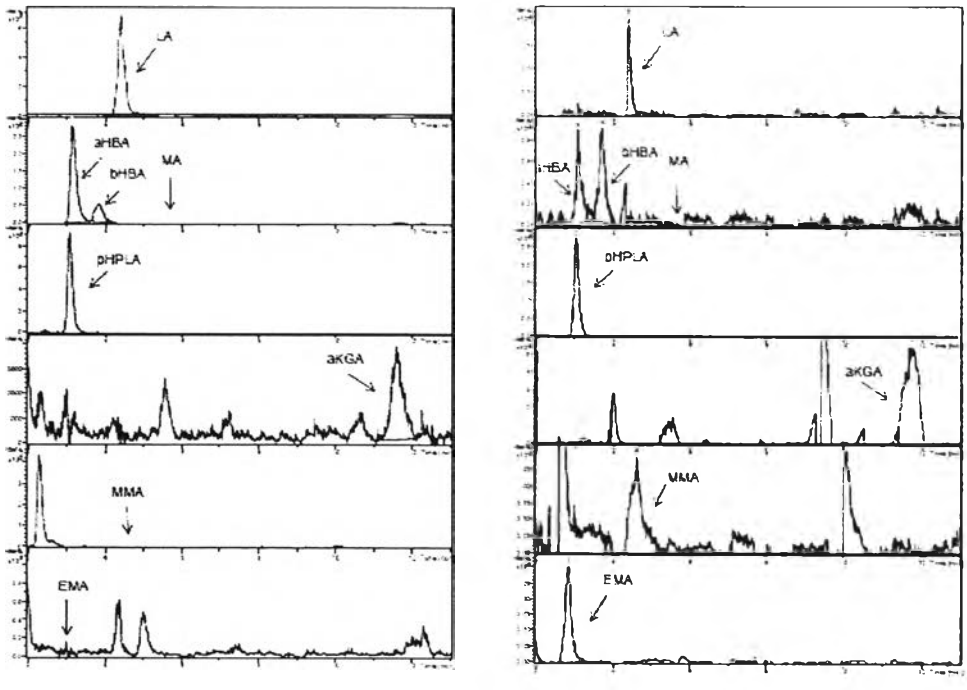

Figure 18 Extracted chromatogram of acids found in a plasma sample (A) and in a urine sample (B) from a patient with severe falciparum malaria.

Measured concentrations in Figure 3A: LA= 1169 µg/mL, aHBA= 38.4 µg/mL, bHBA= 38.0 µg/mL, pHPLA= 4.30 µg/mL and aKGA= 32 µg/mL.

Measured concentrations in Figure 3B: LA= 513 µg/mL, aHBA= 69.5 µg/mL, bHBA= 111 µg/mL, pHPLA= 27.7 µg/mL, MMA= 13.3 µg/mL, EMA= 48.1 µg/mL and aKGA= 107 µg/mL [54].

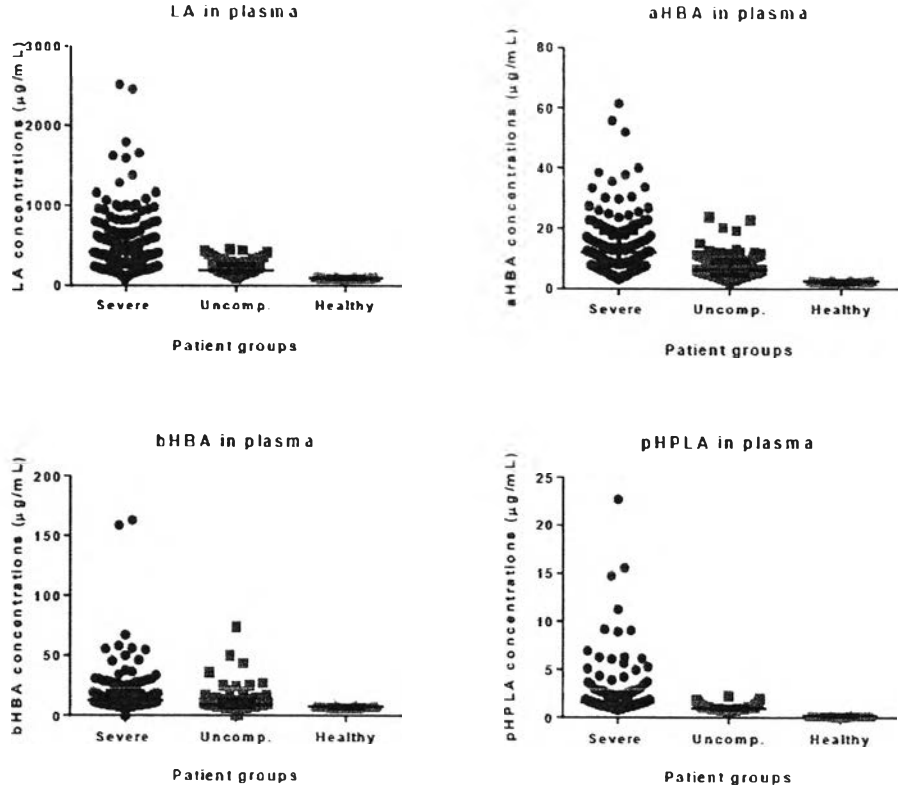

Figure 19 Scatter plots of all 296 plasma samples (3 groups; severe malaria (n=141), uncomplicated malaria (n=87) and healthy control (n=68)) with 4 acids (LA, aHBA, bHBA and pHPLA)

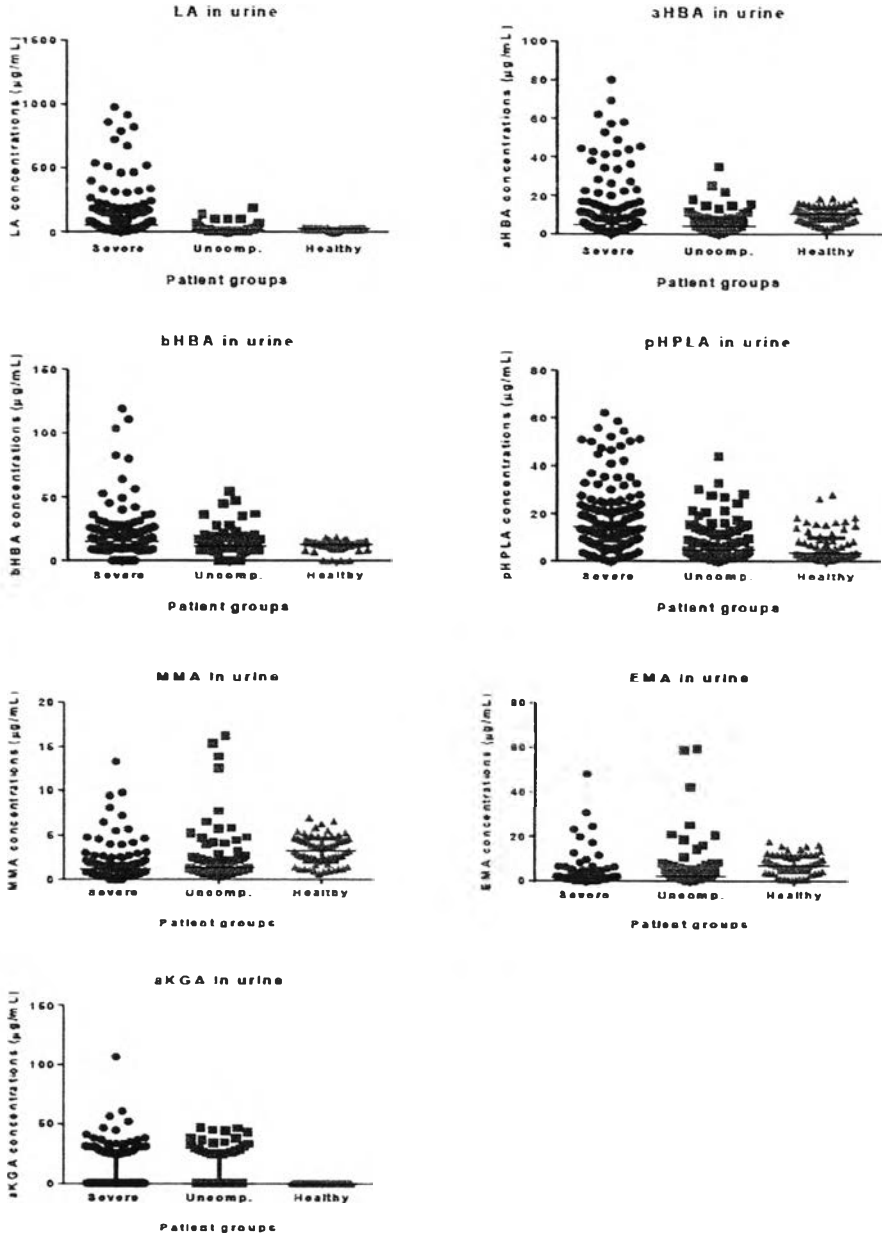

Figure 20 Scatter plots of all 288 urine samples (3 groups; severe malaria (n=133), uncomplicated malaria (n=94) and healthy control (n=61)) with 7 acids (LA, aHBA, bHBA, pHPLA, MMA, EMA and aKGA)

4.3 Statistical Data Analysis

4.3.1 Univariate Analysis

Concentration of four acids (LA, aHBA, bHBA and pHPLA) and seven acids (LA, aHBA, bHBA, pHPLA, MMA, EMA and aKGA) from three groups of plasma and urine samples (severe malaria, uncomplicated malaria and healthy controls) were tested with Kruskal-Wallis nonparametric test since the distribution of all acids in plasma and urine samples were non-normal. The multiple comparison test of median for each of all acid concentrations between 2 groups (severe vs. uncomplicated, severe vs. healthy and uncomplicated vs. healthy) was performed individually.

The statistical test results showed that median of all four acids (LA, aHBA, bHBA and pHPLA) in plasma (Figure 19 and Table 7) were different significantly between three sample groups with p-value < 0.0001. The multiple comparison test between 2 groups also showed that they were all different significantly with p-value < 0.0001. The median concentration of all acids in plasma showed that in severe malaria group had the highest concentration, the second high was uncomplicated group and healthy had the lowest concentration (Table 7). Especially, pHPLA in severe group had median concentration more than eight times compares with healthy group. Furthermore, LA and aHBA in severe group had median concentration more than four times compares with healthy group. Figure 19 showed that the scatter of four acids in severe malaria group had a large variation compare with the other groups that might be resulted from the broad clinical criteria for identifying sample group (see appendix) and genetic variation of patients. The scale of concentration among four acids were huge different, especially in LA was more than thirty time of concentration level compare with other three acids. Therefore, data scaling (in section 4.3.2.1) was nescessary very much before processing multivariate data analysis for making all acid concentrations in comparable scale. Furthermore, the statistical results presented above could be concluded that no acid could be excluded to classify three sample groups of patients since all four acids were important and could be potential contributors of acidosis in severe malaria. To consider that all four acids were important equally or not, multivariate data analysis was needed to perform for considering them simultaneously.

In addition, the results of urine samples presented that median concentration of five acids (Figure 20 and Table 8) were different significantly between three sample groups with p-value < 0.0001. However, bHBA and pHPLA were not different significantly between three sample groups with p-value < 0.0001. The multiple comparison tests between 2 groups also showed the same results in all seven acids with p-value < 0.0001. The scatter plot of all seven acid concentrations (Figure 20) showed that there was a huge variation in all three sample groups (severe, uncomplicated and healthy) since there were many variation factors effect acid concentration in urine (i.e. volume of water that patients consume before collecting samples and might be different ability of patients to excrete acids). Therefore, urine acid concentration profiles were considered as additional information in this study. The median concentration trend for seven acids in urine was not similar to in plasma. aHBA, MMA and EMA showed median concentration of healthy group was higher than patient groups (severe and uncomplicated malaria). For LA and bHBA, median concentration of severe group was the highest but healthy group has concentration higher than uncomplicated group. Only pHPLA median concentration in urine had the same trend as in plasma that severe malaria group had the highest concentration, the second high was uncomplicated group and healthy had the lowest concentration. aKGA median concentration of all three groups were lower than method limit of detection (LOD). From the results of median concentration profiles presented above, it was difficult to consider that which acids was important and could be excluded or not as contributors of acidosis in severe malaria. Therefore, multivariate data analysis was also nescessary for considering all acid concentration profiles simultaneously.

Table 7 Medians with interquartile range of four acid concentrations in plasma from all three sample groups

Sample groups	Severe	Uncomp.	Healthy	Severe	Uncomp.	Healthy
Number of samples	141	87	68	141	87	68
Acids	LA			aHBA		
Concentration (μg/mL)						
Median	416	193	101	12.2	6.20	2.37
Interquartile range	267-686	158-262	93.9-106	8.27-18.7	4.60-9.33	2.07-2.69

Sample groups	Severe	Uncomp.	Healthy	Severe	Uncomp.	Healthy
Number of samples	141	87	68	141	87	68
Acids	bHBA			pHPLA		
Concentration (μg/mL)						
Median	13.0	8.92	7.45	1.63	1.05	0.200
Interquartile range	9.76-20.1	7.37-12.4	6.65-7.93	1.22-2.44	0.910-1.16	0.160-0.230

Severe: Severe malaria; Uncomp: Uncomplicated malaria; Healthy: Healthy volunteers; LA: L-lactic acid; aHBA: α-hydroxybutyric acid; bHBA: β-hydroxybutyric acid; pHPLA: p-hydroxyphenyllactic acid

Table 8 Medians with interquartile range of seven acid concentrations in urine from all three sample groups

Sample groups	Severe	Uncomp.	Healthy	Severe	Uncomp.	Healthy
Number of samples	133	94	61	133	94	61
Acids	LA			aHBA		
Concentration (µg/mL)						
Median	56.7	13.2	24.4	4.95	3.62	10.4
Interquartile range	19.5-156	8.28-23.9	22.1-27.4	2.23-13.9	1.97-7.47	7.51-13.3

Sample groups	Severe	Uncomp.	Healthy	Severe	Uncomp.	Healthy
Number of samples	133	94	61	133	94	61
Acids	bHBA			pHPLA		
Concentration (µg/mL)						
Median	15.5	11.1	13.2	14.4	5.06	3.66
Interquartile range	9.53-23.9	8.17-15.8	10.2-14.5	8.29-24.3	1.92-12.7	2.04-10.1

Sample groups	Severe	Uncomp.	Healthy	Severe	Uncomp.	Healthy
Number of samples	133	94	61	133	94	61
Acids	MMA			EMA		
Concentration (μg/mL)						
Median	1.09	1.39	3.25	1.61	1.99	7.09
Interquartile range	0.425-1.78	0.855-2.35	2.09-4.64	0.655-3.51	1.31-4.67	4.33-11.6

Sample groups	Severe	Uncomp.	Healthy
Number of samples	133	94	61
Acids	aKGA		
Concentration (μg/mL)			
Median	<LOD	<LOD	<LOD
Interquartile range	<LOD-27.9	<LOD-25.1	<LOD

Severe: Severe malaria; Uncomp: Uncomplicated malaria; Healthy: Healthy volunteers; LA: L-lactic acid; aHBA: α-hydroxybutyric acid; bHBA: β-hydroxybutyric acid; pHPLA: p-hydroxyphenyllactic acid

4.3.2 Multivariate Analysis

4.3.2.1 Data Pre-processing

In this study, acid concentration profiles in plasma samples were considered as the main contributors of acidosis in severe malaria thus they were evaluated primarily in this section. However, acid concentration profiles in urine were also considered as just additional information according to reasons explained in section 4.3.1.

The main objectives of data pre-processing step were to transform quantitative analytical raw data into data matrix which was simple format before processing in multivariate data analysis (pattern recognition step). In addition, the other aims were to adjust all data in the same scale for comparing and to reduce variation (i.e. different experimental date, instrument condition, etc.) between all samples. Therefore, data matrix from 296 plasma samples with 4 acid concentration profiles had 296 rows and 4 columns (296X4) and also from 288 urine samples with 6 acid concentration profiles gave data matrix (288 X 6). Both data matrix from plasma and urine were scaled by transforming the elements step, row and column scaling step, respectively.

In transforming the element step, nth root technique was chosen since it could be used with zero values (logarithmic scaling method could not be use) and could perform nth root for further reducing of different scale. Row scaling step was then used to make data of samples comparable. According to variation of analyzed compound concentrations from sample to sample, this step was performed by scaling all elements of each sample to the quantities of an internal standard (IS).
In the last step (column scaling), acid concentration scales among four acids in plasma samples were huge different (especially in LA compare with other acids). To ensure that all significant variables (concentration profiles of all acids) could have a similar influence on the results (the result was not dominated by the acid which has highest concentration level), the column scaling techniques (mean centering and standardization) were tested and compared [13,[55, 56]].

Figure 21 showed comparison of the classification feature for all three groups of plasma samples presented as 2D score plot from PCA technique. Due to the primary objective of this study that was to identify and classify healthy control group vs. malaria patient groups (which including severe and uncomplicated

malaria). Last Sub Figure in Figure 21, shows the best classification achieved by using n^{th} root method as a transforming the elements step and standardization method as data scaling step. Therefore, these two techniques were chosen to perform in data pre-processing step and applied to all data matrix of plasma and urine samples before processing pattern recognition steps (section 4.3.2.2 and 4.3.2.3).

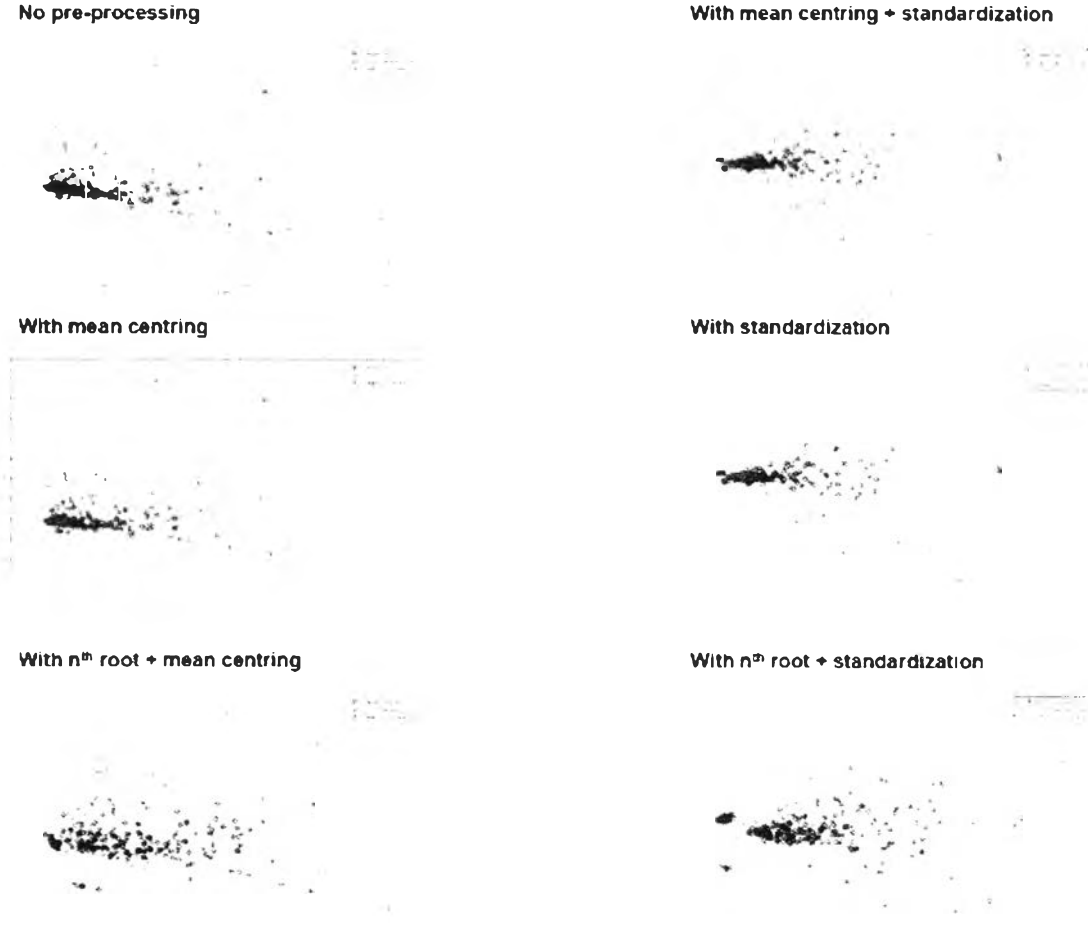

Figure 21 Two dimension score plots (from PCA technique) of all 3 groups of plasma samples (n=296) with 4 targeted acid variables using different pre-processing methods

(Blue color = severe malaria, red color = uncomplicated malaria and green color = healthy)

4.3.2.2 Unsupervised Pattern Recognition

Unsupervised pattern recognition is commonly initial method used for exploration and visualization the information underlying data without any prior knowledge of sample class. Principal component analysis (PCA) is the most frequently used technique for finding possible trend in the samples and the discriminant power between the variables. Reduction of the data matrix dimension to a very small number of orthogonal principal components (usually 2 or 3 PCs) is an approach to summarize all variances in the original data matrix. These new latent variables (PCs) are linear combinations of the variables which explain the greatest variance on PC1 and the next greatest variance on PC2 and PC3, respectively. PCA decomposes the data into score vectors represent for samples and loading vectors represent for variables (targeted acids). [13,[57]].

The two and three dimension score plots of plasma samples with considering four targeted acid variables from three sample groups (Figure 22) showed that healthy control group could classify from two malaria patient groups completely with variance of three first PCs (73.1, 15.4 and 7.8%, respectively). However, severe malaria group could not classify from uncomplicated malaria group completely as there were some areas of these 2 groups overlapping. For an evaluation of the discriminant power between the variables (candidate organic acids), Figure 23 (bi-plot) showed relationships between samples and variables that all acids have tendency to be increased in severe malaria patient group. In addition, loadings of the variables in the first PC (Table 9) showed that all variables (acids) have nearly equal discriminant power. Therefore, according to factor loading results, all 4 acids (LA, aHBA, bHBA and pHPLA) were considered equally to be potential for acidosis in severe malaria. These results allowed the initial observation for classification of all samples.

Figure 24 (PCA bi-plot) showed the relationships between urine samples in the three groups and also between urine samples and seven acid variables. The results showed that when considering seven acids simultaneously, healthy group could not be completely classified from two malaria patient groups. There were overlapping areas between three sample groups with variance of two first PCs (37.1 and 22.4%, respectively). Table 10 presents loadings of the variables in the first PC, the results show that LA, aHBA and bHBA have more discriminant power than the other four acids (pHPLA, MMA, EMA and aKGA). However, Figure 25 (PCA bi-plot) which considered only LA, aHBA and bHBA showed that healthy group

still could not be classified from two malaria patient groups completely since there were overlapping areas between three sample groups.

Therefore, only plasma acid concentration profile could be used as the classifier and the potential contributor to acidosis in severe malaria. Subsequently, plasma acid concentration profile would be processed by supervised pattern recognition in the next step (section 4.3.2.3) to construct predicted model for a classification of two groups of malaria patient and healthy group (the primary objective) and also three groups of severe malaria, uncomplicated malaria and healthy group as the secondary objective.

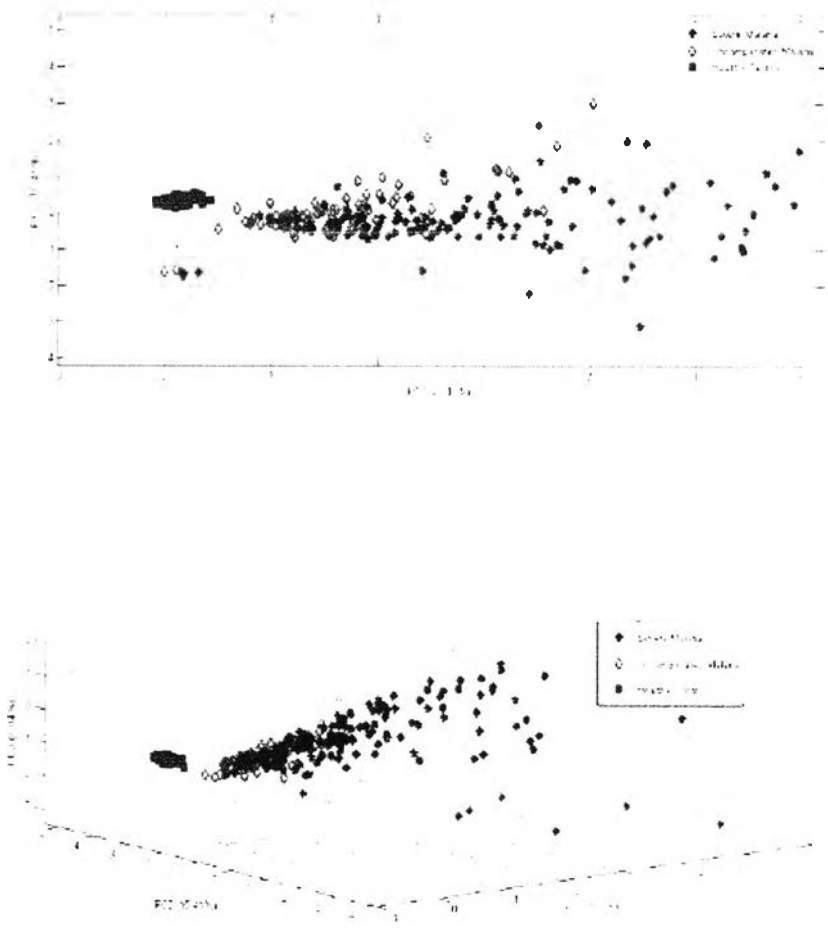

Figure 22 Two dimension score plot of plasma samples (n=296) with 4 targeted acid variables from three sample groups (A) and Three dimension score plot (B)

(Blue color = severe malaria, red color = uncomplicated malaria and green color = healthy)

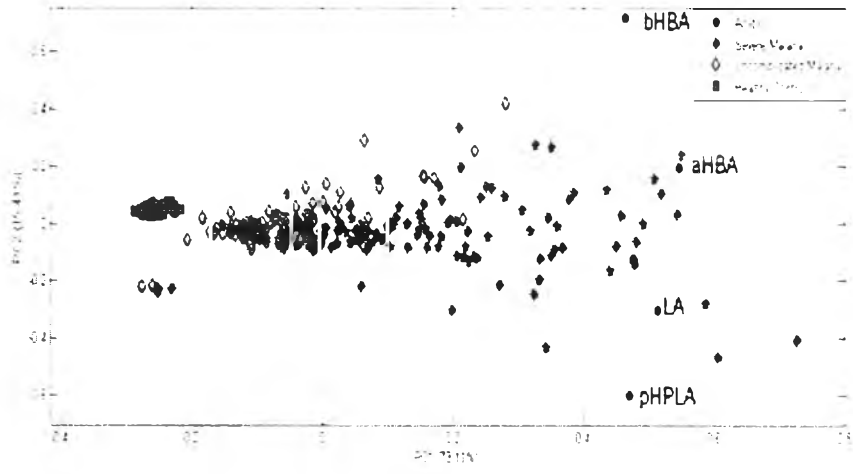

Figure 23 Bi-plot (2D) of plasma samples (n=296) with 4 targeted acid variables from three sample groups

(Blue color = severe malaria, red color = uncomplicated malaria and green color = healthy)

Table 9 Factor loadings in the two first principal components (PC) of data matrix from plasma samples

Variable	PC 1	PC 2
LA	0.5138	-0.3044
aHBA	0.5471	0.1942
bHBA	0.4641	0.7151
pHPLA	0.4705	-0.5986

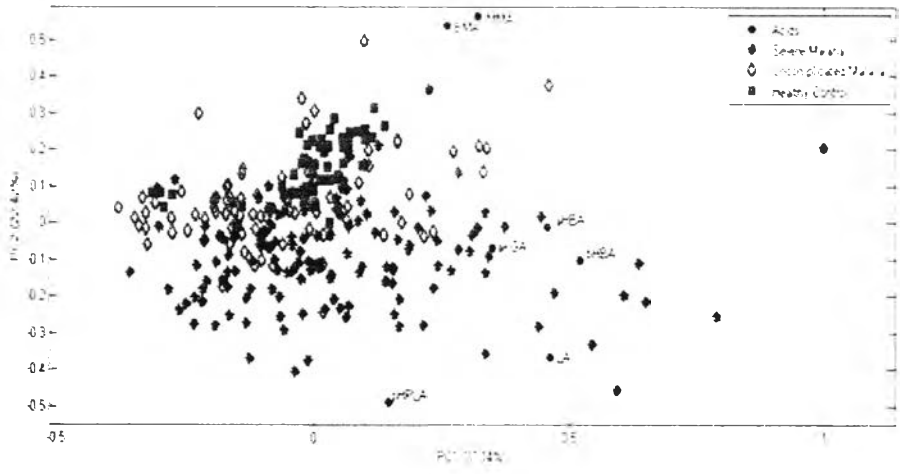

Figure 24 Bi-plot (2D) of urine samples (n=288) with 7 targeted acid variables from three sample groups

(Blue color = severe malaria, red color = uncomplicated malaria and green color = healthy)

Table 10 Factor loadings in the two first principal components (PC) of data matrix from urine samples

Variable	PC 1	PC 2
LA	0.4606	-0.3657
aHBA	0.4559	-0.0114
bHBA	0.5197	-0.1022
pHPLA	0.1436	-0.4909
MMA	0.3185	0.5641
EMA	0.2597	0.5405
aKGA	0.3469	-0.0657

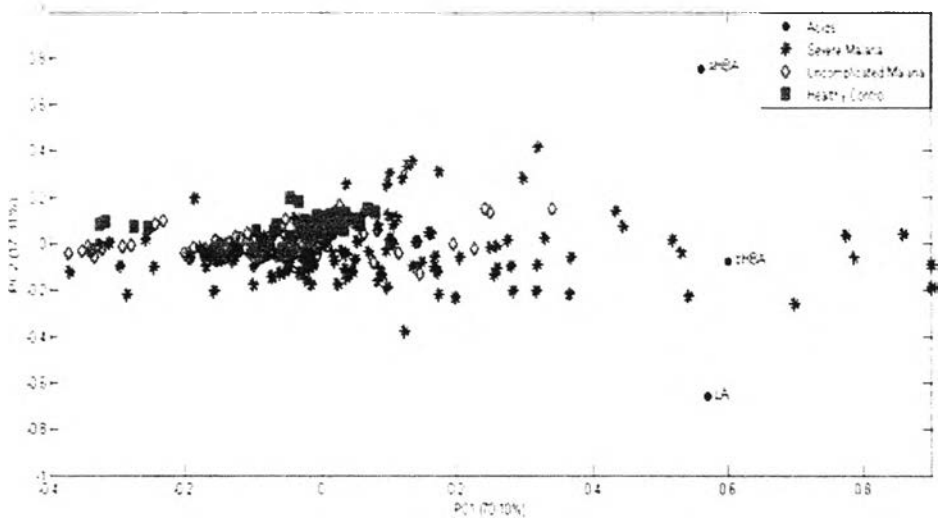

Figure 25 Bi-plot (2D) of urine samples (n=288) with 3 targeted acid variables from three sample groups

(Blue color = severe malaria, red color = uncomplicated malaria and green color = healthy)

Figure 23 (PCA) bi-plots of plasma showed that there was huge variation in severe group (blue color) and uncomplicated group (red color), these might be resulted from the broad clinical criteria for identifying sample group (see appendix) and genetic variation of patients. Figure 24 and Figure 25 shows that acid concentration profiles in urine samples (consider seven and three acids, respectively) could not be used to classify any of three sample groups (severe, uncomplicated and healthy) since there were vast variation in all of three groups that might be caused by variation from much more factors compare with in plasma (i.e. volume of water that patient consume, different ability to excrete acids of patients, etc.).

4.3.2.3 Supervised Pattern Recognition

Supervised pattern recognition is a technique usually used to characterize differences among samples (i.e. disease and healthy class) and confirm result from unsupervised pattern recognition (i.e. PCA). This technique uses knowledge of sample class to build the predicted model for classification of new unknown samples to one of known class based on its pattern of measurements. The

most potential advantage of this technique is its ability to create the model by using knowledge sample class only once in order to predict class of unlimited unknown samples [13-14, [58]].

In this study, PLSDA and LDA were applied to construct the predicted model for the classification of two sample groups (malaria patient and healthy) and three sample groups (severe malaria, uncomplicated malaria and healthy), respectively. The most important step of analyzing supervised pattern recognition results is cross-validation to determine the predictive power of developed model. Predictive power (express as sensitivity and specificity) is a performance efficiency of developed model to predict class of the unknown samples. Leave-one-out cross validation was chosen to validate PLSDA and LDA predicted model, since it is simple, effective and suitable technique for validation of large sample numbers [13].

Partial least square analysis (PLSDA) is one of the most common regression and supervised linear modelling technique. PLSDA has a similar principal as PCA which decomposes the data into score vectors represent for samples and loading vectors represent for variables (targeted acids), however PLSDA considers both experimental data and class vector while PCA considers only experimental data without prior class knowledge [13-14]. The advantage of PLSDA is flexibility, particularly in the situation where the number of variables are very much more than the number of samples. The discriminatory power of PLSDA is similar to LDA if appropriate data scaling is applied. PLSDA is a high effective technique to classify two groups of samples [13]. Therefore, it was used for the primary objective to classify and predict malaria patient group (including severe and uncomplicated malaria) and healthy group in this study. Table 11 is a contingency table presents classification result of PLSDA for total plasma samples (n=296); including malaria patient group (n=228) and healthy group (n=68). The results (Table 11) indicate good sensitivity (91.7%) and excellent specificity (100.0%) in cross validated prediction for classifying malaria patient group vs. healthy control group. These results are proved that PLSDA predicted model of this study gave a reliable classification and predictive ability.

Table 11 Summary of PLSDA classification result

Predicted Class	True class	
	Patients	Healthy
Patients	209	0
Healthy	19	68
Total	228	68

LDA is commonly used to calculate discriminant functions as linear combinations of selected variables for separating the groups [13]. LDA method has the potential to classify samples from more than two groups and with a discriminatory power similar to PLSDA in case of appropriate data scaling. Furthermore, data scaling (pre-processing step) was applied with the same techniques for both PLSDA and LDA in this study. Therefore, LDA was chosen for secondary objective to classify three sample groups (severe malaria, uncomplicated malaria and healthy group). Table 12 is a contingency table presents classification result of LDA for total plasma samples (n=296); including severe malaria (n=141), uncomplicated malaria (n=87) and healthy group (n=68). Sensitivity and specificity of all three groups are summarized in Table 13.

Performance indices of LDA model (Table 13) showed that the classification of healthy control group is perfect (100% sensitivity), uncomplicated malaria group is good (82.8% sensitivity). However, severe malaria group shows just fair classification with a sensitivity of only 64.5%.

From Table 11 to Table 13 and results presented above showed that four acid concentration profiles in plasma samples could be used for completely classification and achievement of predicted model with excellent results for malaria patient group vs. healthy group. In addition, if specific only severe malaria group vs. healthy group, the classification and prediction results were excellent the same. However, these four acid concentration profiles in plasma samples could not use to classify severe group vs. uncomplicated group completely since there were vast variation of samples in these two groups (might be result from genetic variation

of patients and the broad clinical criteria for identifying malaria patient sample group (see appendix)).

Table 12 Summary of LDA classification result

	True Class		
Predicted Class	Severe malaria	Uncomp. Malaria	Healthy control
Severe malaria	91	7	0
Uncomp. Malaria	49	72	0
Healthy control	1	8	68
Total	141	87	68

Table 13 Sensitivity and specificity of LDA in three sample groups

Class	Sensitivity (%)	Specificity (%)
Severe malaria	64.5	95.2
Uncomp. malaria	82.8	76.4
Healthy control	100.0	94.8