

# BMJ Open The association of tryptophan and phenylalanine are associated with arsenic-induced skin lesions in a Chinese population chronically exposed to arsenic via drinking water: a case-control study

Yaping Wei <sup>1,2,3</sup>, Chaonan Jia,<sup>1,2</sup> Yuan Lan,<sup>4</sup> Xiangqing Hou,<sup>1,2</sup> Jingjing Zuo,<sup>4</sup> Jushuang Li,<sup>1,2</sup> Tao Wang,<sup>1,2</sup> Guangyun Mao <sup>1,2,5</sup>

**To cite:** Wei Y, Jia C, Lan Y, *et al.* The association of tryptophan and phenylalanine are associated with arsenic-induced skin lesions in a Chinese population chronically exposed to arsenic via drinking water: a case-control study. *BMJ Open* 2019;**9**:e025336. doi:10.1136/bmjopen-2018-025336

► Prepublication history and additional material for this paper are available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2018-025336>).

YW and CJ contributed equally.

Received 13 July 2018

Revised 03 September 2019

Accepted 01 October 2019



© Author(s) (or their employer(s)) 2019. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

## Correspondence to

Professor Guangyun Mao; [maogy@wmu.edu.cn](mailto:maogy@wmu.edu.cn)

## ABSTRACT

**Objectives** We investigated the association of specific serum amino acids (AAs) with the odds of arsenic-induced skin lesions (AISL) and their ability to distinguish patients with AISL from people chronically exposed to arsenic.

**Design** Case-control study.

**Setting** Three arsenic-exposed villages in Wuyuan County, Hetao Plain, Inner Mongolia, China were evaluated.

**Participants** Among the 450 residents aged 18–79 years, who were chronically exposed to arsenic via drinking water, 56 were diagnosed as having AISL (defined as cases). Another 56 participants without AISL, matched by gender and age ( $\pm 1$  year) from the same population, were examined as controls.

**Main outcome measures and methods** AA levels were determined by ultra-high-performance liquid chromatography–quadrupole time-of-flight mass spectrometry-based metabolomics analysis. Potential confounding variables were identified via a standardised questionnaire and clinical examination. Multivariable conditional logistic regression model and receiver operating characteristic curve analyses were performed to investigate the relationship between specific AAs and AISL.

**Results** Tryptophan and phenylalanine levels were negatively associated with AISL ( $p < 0.05$ ). Compared with that in the first quartile, the adjusted OR of AISL in the second, third and fourth quartiles were decreased by 44%, 88% and 79% for tryptophan and 30%, 80% and 80% for phenylalanine, respectively. The combination of these two higher-level AAs showed the lowest OR for AISL (OR=0.08; 95% CI 0.02 to 0.25;  $p < 0.001$ ). Furthermore, both AAs showed a moderate ability to distinguish patients with AISL from the control, with the area under the curve (AUC; 95% CI) as 0.67 (0.57 to 0.77) for tryptophan and 0.70 (0.60 to 0.80) for phenylalanine ( $p < 0.05$ ). The combined pattern with AUC (95% CI) was 0.72 (0.62 to 0.81), showing a sensitivity of 76.79% and specificity of 58.93% ( $p < 0.001$ ).

**Conclusions** Specific AAs may be linked to AISL and play important roles in early AISL identification.

**Trial Registration number** NCT02235948.

## Strengths and limitations of this study

- Our findings were based on a community-based metabolomics study with paired design, and strict quality assurance and quality control.
- Multivariable conditional logistic models were used to examine the association between specific levels of amino acids (AAs) and arsenic-induced skin lesions (AISL), and the receiver operator characteristic analysis was applied to evaluate the feasibility of AAs to distinguish patients with AISL from their counterparts.
- Although AAs were determined by untargeted metabolomics approach, which can assess a large number of metabolites precisely and efficiently, only relative levels of AAs could be obtained instead of their accurate quantitative concentration.
- Based on a case-control study, the findings only revealed the association between AAs and the odds of AISL rather than confirming their causal relationship.
- The participants were mainly chronically exposed to arsenic via drinking water, which may limit the findings extrapolated to another arsenic exposure population via food or other routes.

## INTRODUCTION

Chronic arsenic exposure via drinking water is widely considered a global health concern affecting several people worldwide. This exposure may cause various human health issues such as cardiovascular diseases, diabetes and cancer.<sup>1 2</sup> With the industrial boom and considerable increase in global water pollution including arsenic contamination in the past, the prevalence and burden of arsenic-induced health damage will continue to increase. The skin has been confirmed as one of the most common and susceptible targets of arsenic-induced health effects.

Cutaneous skin lesions are typical signs of arsenicosis after persistent and long-term arsenic exposure. These lesions are characterised by hyperkeratosis and hyperpigmentation. Considerable evidence of the prevalence of arsenic-induced skin lesions (AISL) has been reported in several countries.<sup>3–5</sup>

As AISL are widely accepted as the major early manifestation of arsenic toxicity<sup>6</sup> and may be an indicator of susceptibility to more serious arsenic-induced health hazards,<sup>7</sup> it is critical to identify people who are at risk as early as possible to prevent the onset or delay the progression of serious health problems. Several possible mechanisms such as genetic differences,<sup>8</sup> oxidative stress<sup>9</sup> and epigenetic dysregulation<sup>10</sup> may explain arsenic poisoning. Previous studies also suggested that arsenic methylation in vivo is associated with metabolic syndrome.<sup>11 12</sup>

Amino acids (AAs) are the 'basic unit' of all proteins and are necessary to maintain health. Some AAs are important regulators of key metabolic pathways and also help in maximising food utilisation, enhancing protein accretion and improving health.<sup>13 14</sup> Abnormal metabolism of AAs can disturb homeostasis in the body, impair growth and development and even cause death.<sup>15</sup> Thus, the levels of serum AAs may be important indicators of metabolic status and disease condition. As powerful tools in system biology research, metabolomics approaches are beneficial for unbiased monitoring of changes in endogenous metabolism-related physiological processes, providing integrative information on distinct features across multiple functional levels. These methods capture the core attributes responsible for various phenotypes, which are particularly important in understanding the relevant pathophysiological changes of a disease and its status, and in identifying novel biomarkers for risk screening, diagnosis, treatment and prognosis of important human diseases.<sup>16–18</sup>

Animal experiments and epidemiological studies have reported obvious arsenic-related metabolomics perturbations.<sup>19 20</sup> These results suggest that the relationship between specific metabolites and arsenic-induced health lesions should be investigated. However, only a few studies have been conducted to comprehensively examine the metabolic mechanism relevant to AISL, particularly for AA metabolism. This study was conducted to quantitatively examine the association of several specific AAs with AISL and their ability to identify AISL.

## METHODS

### Study population

The study data were originally obtained from a randomised, double-blind and placebo-controlled clinical trial performed in 2010, in which all subjects were randomly selected by permuted block randomisation from a single rural area in which a population was chronically exposed to low-level arsenic drinking water, had similar lifestyles and were influenced by similar environmental factors. Information on the inclusion and exclusion

criteria of the participants can be found in our previous study.<sup>21</sup> Strictly following the criteria of arsenicosis,<sup>22</sup> AISL was diagnosed as the presence of arsenic-induced keratosis, hyperpigmentation or depigmentation by a physician from the Wenzhou Medical University at the beginning of the trial. This was a matched case–control study (1:1 matching). Among 450 residents aged 18–79 years enrolled in the previous trial, 56 were diagnosed as having AISL and selected as the case group. Another 56 participants without AISL matched by gender and age ( $\pm 1$  year) from the same population were evaluated as controls. The inclusion criteria were subjects who underwent a metabolomic test. Unmatched participants and those without serum metabolites data were excluded.

### Data collection and assessment

Information on age, gender, exposure year, body mass index, smoking, alcohol consumption and education level, among other factors, was collected using a standardised questionnaire. Blood and urine samples were also collected at the time of participants' enrolment. The detailed methods for analysing blood and urine samples and assessment methods for clinical variables including fasting plasma glucose (FPG), serum urea nitrogen, serum folate, total homocysteine, total cholesterol, triglycerides (TG), high-density lipoprotein, low-density lipoprotein (LDL) and others have been published previously.<sup>21</sup> Various urinary arsenic species were separated and detected by high-performance liquid chromatography coupled mass spectrometry system.<sup>23</sup> The species of arsenic in urine samples consisted of inorganic arsenic (iAs, [iAs<sup>III</sup> plus iAs<sup>V</sup>]), monomethyl arsenate (MMA, [MMA<sup>III</sup> plus MMA<sup>V</sup>]) and dimethyl arsenate (DMA, [DMA<sup>III</sup> plus DMA<sup>V</sup>]). All arsenic species were corrected by creatinine. The total arsenic (tAs) was the sum of iAs, MMA and DMA. The percentages of arsenic species were defined as follows: iAs%=iAs/tAs\*100%, MMA%=MMA/tAs\*100% and DMA%=DMA/tAs\*100%, respectively.

### Ultra-high-performance liquid chromatography–quadrupole time-of-flight mass spectrometry metabolomic profiling

Serum samples were thawed to 4°C, and then 600  $\mu$ L of a mixture of 90% acetonitrile and 10% water was added to each sample (200  $\mu$ L in microcentrifuge tubes). The samples were vigorously mixed for 20 s and then centrifuged for 5 min at 12 000 g (20°C). Four hundred microlitres of each supernatant were transferred to a new tube and dried in a vacuum concentrator centrifuge. The dried samples were re-suspended in 130  $\mu$ L of water (containing 15% acetonitrile), mixed vigorously for 20 s centrifuged as described above. Two microlitres of the supernatant were collected for analysis. Serum metabolic profile was acquired using ACQUITY UPLC/Xevo G2 QTof/MS<sup>E</sup> (Waters Corp., Milford, MA, USA). Chromatographic separation was performed at 50°C using a Waters HSS T3 column (2.1 mm  $\times$  100 mm, 1.7  $\mu$ m; Milford, MA, USA) at a flow rate of 0.4 mL/min. The mobile phase was a

mixture of (A) H<sub>2</sub>O with 0.1% formic acid and (B) methanol with 0.1% as follows: 0 min with 100% A and 0% B, 1 min with 100% A and 0% B, 8 min with 0% A and 100% B, and 13 min with 0% A and 100% B. The mass spectrometer was operated under both electrospray ionisation positive-ion (ESI<sup>+</sup>) mode and negative-ion (ESI<sup>-</sup>) mode. The scan range was 50–1200 m/z. The capillary voltage was set to 3000 and 2500 V, respectively. The desolvation flow rate was 800 L/h at 350°C. Argon was used as the collision gas, and the collision energy was adjusted from 10 to 40 eV in each analysis. Quantum clustering samples were prepared by pooling aliquots of each sample and used to reflect the reliability of further metabolomics analysis. After peak deconvolution, alignment, integration and normalisation, the data including retention time, mass to charge ratio (m/z) and peak intensity were extracted from raw chromatograms using Progenesis QI V.2.0 (Waters Corp.). MS/MS was performed to determine metabolite levels using MarkerLynx Applications Manager V.4.1 (Waters Corp.).

### Distinct metabolite identification

The peak intensities of metabolites for the 56 pairs of subjects were acquired and imported into MetaboAnalyst V.4.0 (<http://www.metaboanalyst.ca/>) for statistical analyses. A partial least squares-discriminant analysis (PLS-DA), which is a supervised and well-accepted pattern recognition approach, was used to differentiate between the cases and controls. The false discovery rate (FDR)-adjusted p value in univariate analysis was determined to reduce the potential effect of false-positive results. The criteria used to select metabolites included variable importance in projection (VIP) scores >1 using PLS-DA and a crude or FDR-adjusted p value <0.05 using Wilcoxon signed-rank test. We identified 70 extracted small molecular metabolites linked to the recognition of AISL. The Human Metabolome Database (<http://www.hmdb.ca>) was used to identify the metabolites, which include four AA metabolites (phenylalanine, tryptophan, leucine and phenylalanylphenylalanine).

### Statistical analysis

The normality of continuous data was assessed using both QQ-plots and Shapiro-Wilk test. The data of cases and controls were analysed using paired t-test if they showed a normal or similar normal distribution. Otherwise, Wilcoxon signed-rank test was used. Differences in the proportion of categorical variables between the two groups were evaluated using McNemar-Bowker test. We first used locally weighted scatterplot smoothing models to estimate the ‘real’ relationship between serum AA levels and the probability of AISL. Next, multivariable conditional logistic regression models were used to examine the association between contributing AA levels and AISL after adjusting for potential confounding factors. The individual effects of AA metabolites on the risk of AISL were quantified separately using the OR and 95% CI as follows: with AAs as a categorical

variable (quartiles) and as a continuous variable (scaled to an IQR). Variables with a p value less than 0.2 in the comparison between two groups were selected as potential confounders. This approach has been widely used in many studies, particularly for small sample size. The variance inflation factor (VIF) was used to examine the potential collinearity among variables. As too many covariates in a multiple regression model can lead to overfitting,<sup>24</sup> we selected no more than four variables as confounding factors to decrease the potential of overfitting when assessing the association between AAs and AISL. Furthermore, as distinct metabolites may be highly related to each other, collinearity should be considered. Thus, we used the VIF based on VIF package of R software to detect potential collinearity among the AAs. A VIF greater than 1.5 is considered to indicate collinearity in the model, and the associated variable is removed. The combined effect of relevant AAs on AISL was also determined using a multivariable logistic regression model. A receiver operator characteristic (ROC) analysis was applied to evaluate the value and feasibility of AAs as potentially sensitive and specific biomarkers for recognising AISL. Data management, analysis and figure drawing were performed using R V.3.4.4 (Copyright 2018 The R Foundation for Statistical Computing). All tests were two-sided and the results with p values <0.05 were considered to indicate significance.

### Patient and public involvement

The present study was designed as an observational study, and as such patients and the public were not involved in the planning, recruitment and conduct of this study. All participants were informed about the purpose of this study and signed informed consent at the beginning of the study. The results of this study have not yet been disseminated to the participants.

## RESULTS

**Table 1** summarises the general characteristics of the study population. Comparisons of the demographic data, clinical features and urinary arsenic species in the 56 pairs of subjects are presented in **table 1**. The median (first and third quartiles) age of AISL population was 50.30 (44.70 and 58.70) for the cases and 50.40 (44.60 and 58.70) years for the controls. Both groups contained the same proportion of females (58.93%), and there was no significant difference in urinary arsenic levels between the two groups. More than half of the subjects had no history of smoking or alcohol consumption. The serum TG level in subjects with AISL was significantly lower than that in the controls (p=0.041). Other variables were similar between the subjects with AISL and controls (p>0.05). This indicates that the participants in the two groups were comparable.

**Table 2** shows that the four AAs, with an FDR-adjusted p value of <0.05 and VIP of >1, were significantly lower in

**Table 1** Demographic characteristics of the study population\*

Variables	AISL (n=56)	Non-AISL (n=56)	P value
Clinical characteristics			
Age (years)	50.30 (44.70, 58.70)	50.40 (44.60, 58.70)	0.425
Exposure year (years)	48.19±11.53	47.62±10.97	0.489
Body mass index (kg/m <sup>2</sup> )	24.12±3.14	23.91±2.86	0.697
Fasting plasma glucose (mmol/L)	4.89 (4.60, 5.25)	5.12 (4.53, 5.40)	0.137
Folate (ng/mL)	4.00 (3.20, 5.10)	4.25 (3.35, 5.40)	0.392
Total homocysteine (µmol/L)	12.30 (10.32, 16.50)	12.67 (11.21, 14.69)	0.961
Blood urea nitrogen (mmol/L)	6.45 (5.42, 7.69)	6.84 (5.36, 8.80)	0.603
Total cholesterol (mmol/L)	4.58 (4.10, 5.69)	4.65 (3.96, 5.95)	0.904
Triglycerides (mmol/L)	1.41 (0.90, 1.74)	1.45 (1.09, 2.29)	0.041
High-density lipoprotein (mmol/L)	1.19±0.34	1.16±0.31	0.675
Low-density lipoprotein (mmol/L)	3.04±0.80	3.25±0.84	0.110
Women (# (%))	33 (58.93)	33 (58.93)	1.000
Cigarette smoking (# (%))	20 (35.71)	22 (39.29)	0.696
Alcohol consumption (# (%))	17 (30.91)	21 (37.50)	0.464
Illiteracy (# (%))	21 (37.50)	15 (25.00)	0.252
Urinary arsenic species†			
iAs%	12.26 (8.13, 14.68)	12.31 (10.04, 16.54)	0.148
MMA%	24.68 (20.11, 29.68)	25.85 (20.90, 31.66)	0.420
DMA%	61.84 (56.62, 71.01)	61.84 (47.85, 64.99)	0.096
tAs (µg/g creatinine)	140.93 (104.41, 208.53)	186.77 (80.11, 217.30)	0.445

\* iAs (iAs<sup>III</sup>+iAs<sup>V</sup>); MMA (MMA<sup>III</sup>+MMA<sup>V</sup>); DMA (DMA<sup>III</sup>+DMA<sup>V</sup>); tAs: As<sup>III</sup>+iAs<sup>V</sup>+MMA+DMA; iAs%=iAs/tAs\*100%; MMA%=MMA/tAs\*100% and DMA%=DMA/tAs\*100%.

† The variables met normal distribution was described with mean±SD; otherwise, median (first quartile, third quartile) was used to describe their features. Number of cases (percentage) was used to describe the proportion of categorical variables between the two groups. AISL, arsenic-induced skin lesions; DMA, dimethyl arsenate; iAs, inorganic arsenic; MMA, monomethyl arsenate; tAs, total arsenic.

the cases than in the controls. Two were aromatic amino acids (AAAs) identified as phenylalanine and tryptophan, one was a branched-chain amino acid (BCAA) identified as leucine, and the last AA was phenylalanylphenylalanine. The individual associations of AAs with AISL are presented in [figure 1](#), which shows obvious ‘dose-response’ relationships.

[Table 3](#) shows that participants in the third and fourth quartiles of the four specific AAs were significantly linked to decreased odds of AISL after adjusting for FPG, LDL, TG and DMA%, as compared with that of their lowest quartiles. The category boundaries of the quartiles are shown in Online supplementary table S1. Significant linear trends were detected between AISL and the four

**Table 2** Distinct metabolites in population with arsenic-induced skin lesions and their counterparts

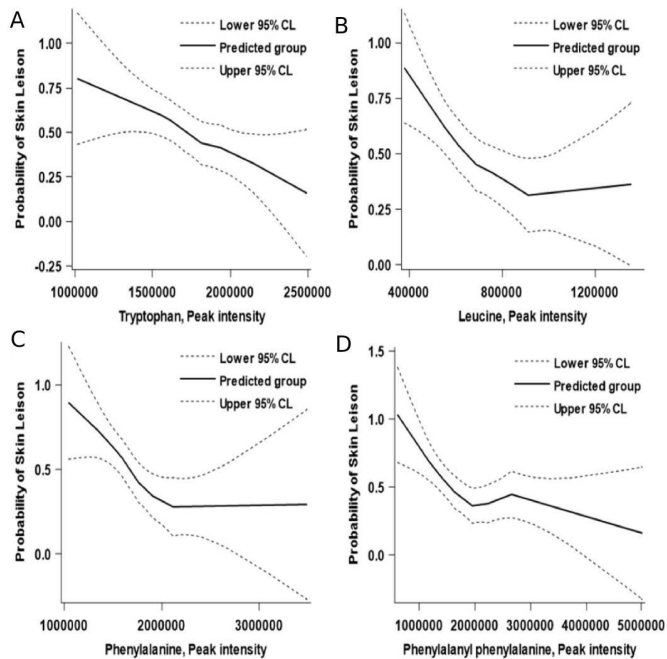
Serum amino acid metabolites	Retention time (min)	Mass-to-charge ratio	VIP value	P values*	Adjusted p values†
Phenylalanine	3.402	166.087	1.508	<0.001	0.009
Tryptophan	3.886	203.082	1.046	0.003	0.014
Leucine	2.642	132.102	1.014	0.001	0.020
Phenylalanylphenylalanine	5.048	313.155	1.833	0.004	0.033

\*Wilcoxon signed-rank test

†Adjusted by false discovery rate.

VIP, variable importance in the projection.





**Figure 1** Association between the peak intensity of tryptophan and phenylalanine and arsenic-induced skin lesions based on multivariable locally weighted regression models. (A) Tryptophan; (B) phenylalanine; (C) leucine and (D) phenylalanylphenylalanine.

serum AAs. The same linear negative association between AISL and per IQR increase in the four serum AAs was observed when these AAs were considered as continuous variables in the present study.

As these four specific AAs were significantly or marginally significantly associated with the odds of AISL, we examined their combined effects on AISL. However, analysis of potential collinearity revealed that among the four specific AAs, both tryptophan and phenylalanine had the smallest VIF value ( $VIF=1.04$ ), and no obvious collinearity existed (Online supplementary table S2). Thus, we mainly focused on tryptophan and phenylalanine to assess the combined effects of AAs on AISL and only presented the results associated with these two AAs in the present study. To avoid the effects of insufficient power because of unreasonable grouping of the results, we classified both tryptophan and phenylalanine into two categories, according to the cut-off values of their mass spectrum peak area based on the ROC analysis, respectively. The higher levels of these two serum AAs were defined as equal to or higher than the cut-off values, whereas the lower categories were considered as less than the associated values.

Table 4 shows the combined effects of tryptophan and phenylalanine levels on AISL after considering the collinearity of variables in the model. The proportions of AISL were 74.3%, 60.0%, 50.0% and 18.2% for participants with lower levels of both tryptophan and phenylalanine (category A), with higher level of tryptophan and lower level of phenylalanine (category B), with lower level of tryptophan and higher level of phenylalanine (category

C), and higher levels of both tryptophan and phenylalanine (category D), respectively. An obvious decrease in the probability of AISL was observed among these four categories. Compared with that of category A, the adjusted OR (95% CI) for participants in categories B, C and D was 0.49 (0.15 to 1.63), 0.32 (0.10 to 1.02) and 0.08 (0.02 to 0.25), respectively. Subjects with higher levels of both tryptophan and phenylalanine had the lowest odds of AISL, which significantly decreased by 92% ( $OR=0.08$ ; 95% CI 0.02 to 0.25;  $p<0.001$ ), after adjusting for the effects of some potential confounding factors. This suggests that tryptophan and phenylalanine are jointly associated with AISL, although no significant interaction between the two AAs and occurrence of AISL was observed ( $p=0.419$ ).

Table 5 shows that, based on the ROC analysis, both serum tryptophan and phenylalanine levels might be potential biomarkers in distinguishing AISL from a chronic arsenic exposure population ( $p=0.0020$  and  $p=0.0017$ ). The area under the curve (AUC) and its related 95% CI, sensitivity, specificity, positive predictive value and negative predictive value were 0.67 (0.57 to 0.77), 69.64%, 62.50%, 65.00% and 67.31% for tryptophan, and 0.70 (0.60 to 0.80), 69.64%, 69.64%, 69.64% and 69.64% for phenylalanine, respectively. The AUC (95% CI), sensitivity, specificity, positive predictive value and negative predictive value of the combination of them were 0.72 (0.62 to 0.81), 76.79%, 58.93%, 65.15% and 71.74% respectively. Our results suggest that these two AAs could be either individually or jointly used as indicators of AISL identification.

## DISCUSSION

In the present study, the association of serum tryptophan and phenylalanine, screened in our previous non-targeted metabolomics study using ultra-high-performance liquid chromatography-tandem mass spectrometer, with AISL and their ability to indicate AISL occurrence were quantitatively evaluated in individual and joint modes. Our results clearly showed that AISL was significantly and negatively associated with serum tryptophan and phenylalanine levels in a chronically arsenic-exposed population via drinking water. Participants with a higher level of both AAs showed the lowest odds of AISL. These two AAs may be useful as indicators of AISL.

The probability of initiation and development of AISL is affected by numerous factors including age, gender, lifestyles, arsenic exposure and metabolism. These factors may be important confounding factors and largely affect our results. To adjust for the effects of these co-factors, we first selected all participants using permuted block randomisation from a single rural area in which the population was chronically exposed to arsenic in the same route, had a similar lifestyle and was influenced by similar environmental factors. Second, the cases and controls were matched by gender and age ( $\pm 1$  year). All of these may be the reason for the so many potential confounders

**Table 3** Relationship of amino acid levels with the odds of arsenic-induced skin lesions\*

Amino acids	N	Cases (%)	Crude		Adjusted†	
			Or (95% CI)	P value	Or (95% CI)	P value
<b>Tryptophan</b>						
Per IQR	112	56(50)	0.48 (0.27 to 0.84)	0.011	0.48 (0.27 to 0.86)	0.013
<b>Quartiles</b>						
Q <sub>1</sub>	28	20 (71.40)	1.00 (1.00 to 1.00)	Ref.	1.00 (1.00 to 1.00)	Ref.
Q <sub>2</sub>	28	15 (53.60)	0.50 (0.16 to 1.54)	0.225	0.56 (0.16 to 1.98)	0.370
Q <sub>3</sub>	28	10 (35.70)	0.12 (0.03 to 0.53)	0.005	0.12 (0.02 to 0.60)	0.010
Q <sub>4</sub>	28	11 (39.30)	0.19 (0.05 to 0.71)	0.014	0.21 (0.05 to 0.84)	0.028
P for trend				0.008		0.012
<b>Phenylalanine</b>						
Per IQR	112	56(50)	0.57 (0.36 to 0.91)	0.019	0.56 (0.33 to 0.94)	0.028
<b>Quartiles</b>						
Q <sub>1</sub>	28	20 (71.40)	1.00 (1.00 to 1.00)	Ref.	1.00 (1.00 to 1.00)	Ref.
Q <sub>2</sub>	28	19 (67.90)	0.79 (0.22 to 2.82)	0.712	0.70 (0.18 to 2.76)	0.609
Q <sub>3</sub>	28	8 (28.60)	0.18 (0.05 to 0.66)	0.010	0.20 (0.05 to 0.77)	0.019
Q <sub>4</sub>	28	9 (32.10)	0.25 (0.08 to 0.79)	0.018	0.20 (0.05 to 0.75)	0.017
P for trend				<0.001		0.001
<b>Leucine</b>						
Per IQR	112	56(50)	0.45 (0.25 to 0.82)	0.019	0.43 (0.21 to 0.86)	0.016
<b>Quartiles</b>						
Q <sub>1</sub>	28	21 (75.00)	1.00 (1.00 to 1.00)	Ref.	1.00 (1.00 to 1.00)	Ref.
Q <sub>2</sub>	28	14 (50.00)	0.33 (0.11 to 1.03)	0.057	0.31 (0.10 to 1.01)	0.052
Q <sub>3</sub>	28	11 (39.30)	0.22 (0.07 to 0.68)	0.009	0.22 (0.07 to 0.73)	0.014
Q <sub>4</sub>	28	10 (35.70)	0.19 (0.06 to 0.59)	0.004	0.19 (0.06 to 0.65)	0.008
P for trend				0.003		0.007
<b>Phenylalanylphenylalanine</b>						
Per IQR	112	56(50)	0.62 (0.36 to 1.04)	0.070	0.71 (0.41 to 1.24)	0.227
<b>Quartiles</b>						
Q <sub>1</sub>	27	21 (77.80)	1.00 (1.00 to 1.00)	Ref.	1.00 (1.00 to 1.00)	Ref.
Q <sub>2</sub>	29	14 (48.30)	0.16 (0.03 to 0.75)	0.021	0.14 (0.03 to 0.73)	0.019
Q <sub>3</sub>	28	9 (32.10)	0.08 (0.02 to 0.39)	0.002	0.09 (0.02 to 0.52)	0.007
Q <sub>4</sub>	28	12 (42.90)	0.12 (0.02 to 0.57)	0.008	0.11 (0.02 to 0.66)	0.016
P for trend				0.006		0.023

\*Adjusted for plasma glucose, low-density lipoprotein, triglyceride and urinary dimethyl arsenate.

†Values are ORs (95% CIs) for arsenic-induced skin lesions from conditional logistic regression. Q<sub>1</sub>: the first quartile; Q<sub>2</sub>: the second quartile; Q<sub>3</sub>: the third quartile; Q<sub>4</sub>: the fourth quartile.

including arsenic exposure, which did not differ significantly between the cases and controls.

Participants enrolled in the current study were chronically exposed to arsenic via drinking water. The geometric mean (GM) and its related 95% CI of urinary iAs/creatinine and tAs/creatinine in this population were 17.49 (14.90 to 20.53) µg/g and 147.20 (129.00 to 167.97) µg/g, respectively. They were much higher than those in the 20 µg/L exposed to arsenic via drinking water (GM (95% CI): 0.4 (0.3 to 0.5) µg/g for iAs and 9.1 (6.5 to 12.7) µg/g for tAs), while obviously lower

than those in the 90 µg/L exposed group (GM (95% CI): 39.4 (31.4 to 49.6) µg/g for iAs and 248.7 (208.8 to 296.3) µg/g for tAs).<sup>25</sup> An available report has shown that AISL cannot be completely cured even though medical technology has already made great progress.<sup>26</sup> Therefore, it is crucial to identify those who are most likely to progress to overt arsenic damages including AISL among people at risk as early as possible. Metabolomics study, which mainly focuses on thoroughly assessing the variation in metabolites possibly linked to disease occurrence and development, has been widely utilised to help

**Table 4** Joint association between tryptophan and phenylalanine levels with arsenic-induced skin lesions

Tryptophan <cut-off value*	Phenylalanine <cut-off value*	N	Cases (%)	Crude		Adjusted†		
				Or (95% CI)	P value	Or (95% CI)	P value	
Yes	Yes	35	26 (74.3)	1.00 (1.00 to 1.00)	Ref.	1.00 (1.00 to 1.00)	Ref.	
No	Yes	20	12 (60.0)	0.52 (0.16 to 1.68)	0.273	0.49 (0.15 to 1.63)	0.244	
Yes	No	24	12 (50.0)	0.35 (0.12 to 1.04)	0.059	0.32 (0.10 to 1.02)	0.053	
No	No	33	6 (18.2)	0.08 (0.02 to 0.25)	<0.001	0.08 (0.02 to 0.25)	<0.001	
Interaction						0.320	0.49 (0.09 to 2.78)	0.419

\*Adjusted for plasma glucose, low-density lipoprotein, triglyceride and urinary dimethyl arsenate.

†Cut-off value was determined using receiver operator characteristic analysis.

understand the pathogenesis of diseases because of its relevance to the phenotypes compared with other omics study.<sup>27</sup> Moreover, mathematical modelling to assess the linkage between small molecular metabolites and arsenic toxicity has advanced.<sup>28</sup> Developing a simple and interpretable modelling approach for the early detection of arsenic-induced lesions is of great theoretical value and realistic meaning,<sup>29</sup> although it might be difficult due to population-specific complexities and effect of some potential unmeasured covariates such as diet and genetic determinants.

Previous studies have reported that gene–gene and gene–environment interactions are involved in arsenicosis through toxicological mechanisms including genomic instability<sup>30</sup> and oxidative stress.<sup>31</sup> Skin hyperpigmentation and palmoplantar hyperkeratosis may be biomarkers for long-term arsenic exposure and useful for identifying differences in metabolites associated with phenotypes. Metabolite analysis may improve the understanding and identification of AISL. An animal study revealed that examining the disruption in AA metabolism on arsenic exposure in rats may be beneficial for understanding arsenic toxicity.<sup>32</sup> In our previous population-based metabolomics study, we found that serum metabolite alterations were significantly related to the risk of arsenic-induced health damages. In the present study, we found that some BCAAs or AAAs were significantly related to AISL occurrence. Several studies across numerous ethnic backgrounds support the use of BCAAs including

leucine, isoleucine, valine and AAA profiles such as those of phenylalanine, tryptophan and tyrosine as biomarkers for identifying metabolic diseases.<sup>27,33</sup> Zhou *et al* reported that arsenic-induced transformed cells had altered metabolite profiles including downregulation of leucine, tryptophan and phenylalanine in the skin lesion group.<sup>34</sup> Consistent with these findings, the levels of two serum AAAs (tryptophan and phenylalanine) were significantly associated with AISL in our study.

Normal metabolism of AAs is necessary for whole-body homeostasis, growth and development, and health.<sup>15</sup> Studies have reported that changes in the availability of AAAs affect cell signalling, gene expression, brain function and neuroendocrine function.<sup>35</sup> Tryptophan, an AA metabolism-related biomarker, is a sensitive and specific indicator of oxidation. Tryptophan metabolism in mammals is a physiological means of preserving immune homeostasis associated with oxidative stress and inflammation.<sup>36,37</sup> Additionally, phenylalanine can be transformed into specific neurotransmitters such as dopamine and epinephrine via the action of related enzymes. Wu *et al*<sup>38</sup> reported that arsenic exposure can lead to neurotransmitter metabolism, which may explain the reduction in phenylalanine. Furthermore, as a peptide-bound phenylalanine, phenylalanylphenylalanine has been reported to affect protein synthesis and secretion,<sup>39</sup> indicating a relationship between endothelium dysfunction and phenylalanine metabolism disorder. The relationship between AA metabolism and AISL remains unclear. The notable

**Table 5** Combination of diagnostic indicators and ROC analysis results\*

Indicators	AUC (95% CI)	Sensitivity, %	Specificity, %	Predict <sup>+</sup> , %	Predict <sup>-</sup> , %	P value
Tryptophan	0.67 (0.57 to 0.77)	69.64	62.50	65.00	67.31	0.002
Phenylalanine	0.70 (0.60 to 0.80)	69.64	69.64	69.64	69.64	0.002
Combined†	0.72 (0.62 to 0.81)	76.79	58.93	65.15	71.74	<0.001

\*The sensitivities, specificity, positive predictive value and negative predictive value were calculated at their best cut-off points; Predict<sup>+</sup>: positive predictive value; Predict<sup>-</sup>: negative predictive value.

†Combined: tryptophan and phenylalanine. The combination is modelled according to the formula  $\beta_1 X_1 + \beta_2 X_2$ , with  $X_j$  denoting the standardised value for the  $j^{\text{th}}$  amino acid and  $\beta_j$  denoting the regression coefficient from the logistic regression model. AUC, area under the curve; ROC, receiver operator characteristic.



alteration of tryptophan and phenylalanine in this study indicates the occurrence of metabolic disorders due to arsenic exposure. These results are also beneficial for understanding the effects of arsenic toxicity and importance of early identification of exposure for delaying the progression of various arsenic-induced lesions, including AISL.

The main strength of this study is that the findings were based on a community-based, long-term arsenic exposure cohort with well-designed quality assurance and quality control throughout the study. However, there were some limitations to this study. First, although the untargeted metabolomics approach can assess a large amount number of metabolites precisely and efficiently, it only provided relative levels of AAs instead of their accurate quantitative concentration. Second, these findings were mainly based on a case-control study, which only reveals the association between AA metabolism and the odds of AISL rather than confirming their causal relationship. Furthermore, the study included 56 AISL cases-matched and 56 non-AISL controls and the sample size might be potentially insufficient. Although metabolomics studies usually have a sample size of no more than 40 cases in each group,<sup>40 41</sup> our study may be under-powered and thus larger studies are needed in the future. Besides, as it is suggested that the ratio of approximately 10–15 observations per predictor in a logistic regression model will produce reasonably stable estimations,<sup>24</sup> we selected only four covariates in the models due to the small sample size and these results need to be confirmed in further studies. Finally, the participants were mainly exposed to arsenic via drinking water, which would limit the findings from being extrapolated to other arsenic exposure populations via food and other routes. Therefore, additional elaborate population-based studies are needed to verify our findings.

In conclusion, specific AAs may be linked to AISL and AA metabolism may play an important role in early AISL identification. Additional studies are needed to confirm our findings.

#### Author affiliations

<sup>1</sup>Department of Preventive Medicine, School of Public Health and Management, Wenzhou Medical University, Wenzhou, China

<sup>2</sup>Center on Evidence-Based Medicine & Clinical Epidemiological Research, School of Public Health and Management, Wenzhou Medical University, Wenzhou, China

<sup>3</sup>Beijing Advanced Innovation Center for Food Nutrition and Human Health, College of Food Science and Nutritional Engineering, Beijing, China

<sup>4</sup>School of Ophthalmology & Optometry, Wenzhou Medical University, Wenzhou, China

<sup>5</sup>Center on Clinical Research, the Affiliated Eye Hospital of Wenzhou Medical University, Wenzhou, China

**Acknowledgements** We gratefully acknowledge all of the study participants of this project. We wish to thank all of the dedication and hard work of the field team at the department of preventive medicine, school of public health and management, Wenzhou medical university. Our thanks must also go to Dr Kendrick Hii, a native English speaker, for his valuable suggestions and edits to the manuscript.

**Contributors** GM and YW designed the study. CJ participated in collecting data. YL and CJ audited the data. YW, Xh, JL and TW conducted the literature search, YW and CJ conducted statistical analysis and interpreted the results. YW and CJ wrote

the first draft of the manuscript. JZ helped with copyediting. GM reviewed the final manuscript and did substantial contributions.

**Funding** This work was supported by the basic public welfare research project of Zhejiang province (LGF19H260011), the initial scientific research fund (KYQD170301) and the major project of the eye hospital of Wenzhou medical university (YNZD201602).

**Competing interests** None declared.

**Patient consent for publication** Not required.

**Ethics approval** Ethical approval was received from the ethics committee of Wenzhou Medical University, Wenzhou, China.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data availability statement** Data are available in a public, open access repository. There are no data in this work. Data are available upon reasonable request. Data may be obtained from a third party and are not publicly available. No data are available. All data relevant to the study are included in the article or uploaded as supplementary information.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

#### ORCID iDs

Yaping Wei <http://orcid.org/0000-0002-4548-7524>

Guangyun Mao <http://orcid.org/0000-0003-0265-5812>

#### REFERENCES

- Mandal BK, Suzuki KT. Arsenic round the world: a review. *Talanta* 2002;58:201–35.
- Smith AH, Steinmaus CM. *Health effects of arsenic and chromium in drinking water: recent human findings. annual review of public health*. Palo Alto: Annual Reviews, 2009: 107–22.
- Guha Mazumder Det *et al*. Arsenic levels in drinking water and the prevalence of skin lesions in West Bengal, India. *Int J Epidemiol* 1998;27:871–7.
- Guo X, Fujino Y, Kaneko S, *et al*. Arsenic contamination of groundwater and prevalence of arsenical dermatosis in the Hetao plain area, inner Mongolia, China. *Mol Cell Biochem* 2001;222:137–40.
- Lindberg A-L, Rahman M, Persson L-A, A-L L, L-A P, *et al*. The risk of arsenic induced skin lesions in Bangladeshi men and women is affected by arsenic metabolism and the age at first exposure. *Toxicol Appl Pharmacol* 2008;230:9–16.
- Tondel M, Rahman M, Magnuson A, *et al*. The relationship of arsenic levels in drinking water and the prevalence rate of skin lesions in Bangladesh. *Environ Health Perspect* 1999;107:727–9.
- Hsu L-I, Chen G-S, Lee C-H, *et al*. Use of arsenic-induced palmoplantar hyperkeratosis and skin cancers to predict risk of subsequent internal malignancy. *Am J Epidemiol* 2013;177:202–12.
- Luo LR, YY L, Gao YH, *et al*. Association between arsenic metabolism gene polymorphisms and arsenic-induced skin lesions in individuals exposed to high-dose inorganic arsenic in northwest China. *Sci Rep* 2018;8.
- Hughes MF, Beck BD, Chen Y, *et al*. Arsenic exposure and toxicology: a historical perspective. *Toxicol Sci* 2011;123:305–32.
- Ren X, McHale CM, Skibola CF, *et al*. An emerging role for epigenetic dysregulation in arsenic toxicity and carcinogenesis. *Environ Health Perspect* 2011;119:11–19.
- Chen J-W, Wang S-L, Wang Y-H, *et al*. Arsenic methylation, GSTO1 polymorphisms, and metabolic syndrome in an arseniasis endemic area of southwestern Taiwan. *Chemosphere* 2012;88:432–8.
- Wang S-L, Chang F-H, Liou S-H, *et al*. Inorganic arsenic exposure and its relation to metabolic syndrome in an industrial area of Taiwan. *Environ Int* 2007;33:805–11.
- Liu R, Li H, Fan WJ, *et al*. Leucine supplementation differentially modulates Branched-Chain amino acid catabolism, mitochondrial function and metabolic profiles at the different stage of insulin resistance in rats on High-Fat diet. *Nutrients* 2017;9:20.
- Platten M, Wick W, Van den Eynde BJ. Tryptophan catabolism in cancer: beyond IDO and tryptophan depletion. *Cancer Res* 2012;72:5435–40.



- 15 Wu G. Amino acids: metabolism, functions, and nutrition. *Amino Acids* 2009;37:1–17.
- 16 Casadei L, Valerio M, Manetti C. Metabolomics: challenges and opportunities in systems biology studies. *Methods Mol Biol* 2018;2018:327–36.
- 17 Johnson CH, Gonzalez FJ. Challenges and opportunities of metabolomics. *J Cell Physiol* 2012;227:2975–81.
- 18 Hui AY, McCarty WJ, Masuda K, et al. A systems biology approach to synovial joint lubrication in health, injury, and disease. *WIREs Syst Biol Med* 2012;4:15–37.
- 19 Lu K, Abo RP, Schlieper KA, et al. Arsenic exposure perturbs the gut microbiome and its metabolic profile in mice: an integrated Metagenomics and metabolomics analysis. *Environ Health Perspect* 2014;122:284–91.
- 20 Zhang J, Shen H, Xu W, et al. Urinary metabolomics revealed arsenic internal dose-related metabolic alterations: a proof-of-concept study in a Chinese male cohort. *Environ Sci Technol* 2014;48:12265–74.
- 21 Guo X, Cui H, Zhang H, et al. Protective effect of folic acid on oxidative DNA damage: a randomized, Double-Blind, and placebo controlled clinical trial. *Medicine*;2015:e1872.
- 22 Standard of diagnosis for endemic arsenism of China (WS/T211–2001). *the Ministry of Health of the People's Republic of China*. Beijing, 2001.
- 23 Chen X, Guo X, He P, et al. Interactive Influence of *N6AMT1* and *As3MT* Genetic Variations on Arsenic Metabolism in the Population of Inner Mongolia, China. *Toxicol Sci* 2017;155:124–34.
- 24 Peduzzi P, Concato J, Kemper E, et al. A simulation study of the number of events per variable in logistic regression analysis. *J Clin Epidemiol* 1996;49:1373–9.
- 25 Sun G, Xu Y, Li X, et al. Urinary arsenic metabolites in children and adults exposed to arsenic in drinking water in inner Mongolia, China. *Environ Health Perspect* 2007;115:648–52.
- 26 Kuo C-C, Moon KA, Wang S-L, et al. The association of arsenic metabolism with cancer, cardiovascular disease, and diabetes: a systematic review of the epidemiological evidence. *Environ Health Perspect* 2017;125:15.
- 27 Wang TJ, Larson MG, Vasan RS, et al. Metabolite profiles and the risk of developing diabetes. *Nat Med* 2011;17:448–53.
- 28 Lawley SD, Cinderella M, Hall MN, et al. Mathematical model insights into arsenic detoxification. *Theor Biol Med Model* 2011;8.
- 29 Juhasz AL, Herde P, Herde C, et al. Validation of the predictive capabilities of the Sbrc-G in vitro assay for estimating arsenic relative bioavailability in contaminated soils. *Environ Sci Technol* 2014;48:12962–9.
- 30 Kibriya MG, Jasmine F, Parvez F, et al. Association between genome-wide copy number variation and arsenic-induced skin lesions: a prospective study. *Environ Health* 2017;16.
- 31 Priori R, Scivo R, Brandt J, et al. Metabolomics in rheumatic diseases: the potential of an emerging methodology for improved patient diagnosis, prognosis, and treatment efficacy. *Autoimmun Rev* 2013;12:1022–30.
- 32 Wang X, Mu X, Zhang J, et al. Serum metabolomics reveals that arsenic exposure disrupted lipid and amino acid metabolism in rats: a step forward in understanding chronic arsenic toxicity. *Metallomics* 2015;7:544–52.
- 33 Weng L, Quinlivan E, Gong Y, et al. Association of branched and aromatic amino acids levels with metabolic syndrome and impaired fasting glucose in hypertensive patients. *Metab Syndr Relat Disord* 2015;13:195–202.10.1089/met.2014.0132
- 34 Zhou Y, Wang Y, Su J, et al. Integration of microRNAome, proteomics and metabolomics to analyze arsenic-induced malignant cell transformation. *Oncotarget* 2017;8:90879–96.
- 35 Lee JC, Chen MJ, Chang CH, et al. Plasma amino acid levels in patients with colorectal cancers and liver cirrhosis with hepatocellular carcinoma. *Hepato-Gastroenterol* 2003;50:1269–73.
- 36 Munn DH, Mellor AL. Indoleamine 2,3 dioxygenase and metabolic control of immune responses. *Trends Immunol* 2013;34:137–43.
- 37 Pawlak K, Domaniewski T, Mysliwiec M, et al. The kynurenines are associated with oxidative stress, inflammation and the prevalence of cardiovascular disease in patients with end-stage renal disease. *Atherosclerosis* 2009;204:309–14.
- 38 L-L W, Gong W, S-P S, et al. Multiple metal exposures and their correlation with monoamine neurotransmitter metabolism in Chinese electroplating workers. *Chemosphere* 2017;182:745–52.
- 39 Zhou MM, Wu YM, Liu HY, et al. Effects of phenylalanine and threonine oligopeptides on milk protein synthesis in cultured bovine mammary epithelial cells. *J Anim Physiol Anim Nutr* 2015;99:215–20.
- 40 Chen L, Cheng C-Y, Choi H, et al. Plasma metabolomic profiling of diabetic retinopathy. *Diabetes* 2016;65:1099–108.
- 41 Piszcz J, Lemancewicz D, Dudzik D, et al. Differences and similarities between LC-MS derived serum fingerprints of patients with B-cell malignancies. *Electrophoresis* 2013;34:2857–64.