Establishment of A Clinical Diagnostic Model for Gouty Arthritis Based on Serum Biochemical Profile: A Prospective Study

Shang Lyu (✉ lupbao@163.com)

Jiangxi University of Traditional Chinese Medicine <https://orcid.org/0000-0003-0650-5571>

Ruowen Ding

Jiangxi University of Traditional Chinese Medicine

Shilin Yang

Jiangxi University of Traditional Chinese Medicine

Wanyuan Chen

Zhejiang Provincial People's Hospital

Yi Rao

Jiangxi University of Traditional Chinese Medicine

Hui Ouyang

Jiangxi University of Traditional Chinese Medicine

Peng Liu

Jiangxi University of Traditional Chinese Medicine

Yulin Feng

Jiangxi University of Traditional Chinese Medicine

Research article

Keywords: gouty arthritis, disease progression, clinical diagnostic model, serum biochemical profile, multiple logistic regression analysis

Posted Date: August 31st, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-67861/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Establishment of A Clinical Diagnostic Model for Gouty Arthritis Based on Serum Biochemical Profile: A Prospective Study

Shang Lyu¹, Ruowen Ding¹, Shilin Yang¹, Wanyuan Chen², Yi Rao¹, Hui OuYang^{1,3}, Peng Liu¹, Yulin Feng^{1,3,*}

¹ National Pharmaceutical Engineering Center for Solid Preparation in Chinese Herbal Medicine, Jiangxi University of Traditional Chinese Medicine, Nanchang 330006, China

² Departments of Pathology, Zhejiang Provincial People's Hospital, People's Hospital of Hangzhou Medical College, Hangzhou 310004, China

³ State Key Laboratory of Innovative Drug and Efficient Energy-Saving Pharmaceutical Equipment, Nanchang 330006, China

* Correspondence: lupbao@163.com(S.L.); fengyulin2003@126.com(Y.-L.F.); Tel: +86-0791-87119651 (S.L.); +86-0791-87119632 (Y.F.)

Abstract: Objective Gouty arthritis is a chronic disease caused by purine metabolism disorder. The progress of the disease is relatively clear, divided into four stages. Clinical judgment of the stage of gouty arthritis is generally based on blood uric acid levels, bursa fluid examination, there are false positive cases. This paper attempts to construct a clinical diagnostic model based on routine blood test data, so as to avoid tedious procedures such as cystic fluid examination and predict the development direction of the disease. **Methods** Serum samples from 579 patients were collected and divided into training set (379 cases) and validation set (200 cases). Through a series of multivariate statistical analysis and correlation analysis, 12 serum biochemical indicators were selected as risk factors for gouty arthritis to obtain the serum biochemical spectrum of gouty arthritis. Based on the biochemical indicators of 379 patients, the clinical diagnosis model of gouty arthritis was established and visualized with R language to predict the probability of patients at different stages, and the stage with the highest probability was used as the diagnosis result, and the evaluation model of 200 patients was used. **Results** The total ROC of the clinical diagnosis model was 0.95, and the Kappa coefficient of the clinical diagnosis model was 0.80. **Conclusion** This model can be applied to clinical prediction and basic research to improve the accuracy of gouty arthritis staging.

Keywords: gouty arthritis, disease progression, clinical diagnostic model, serum biochemical profile, multiple logistic regression analysis

1. Introduction

Gouty arthritis (GA) is a kind of chronic disease caused by purine metabolic disorders, its progression is relatively clear, can be roughly divided into four stages, Hyperuricemia (HUA), acute gouty arthritis (AGA), during the intermittenment period (DIP) and chronic gouty arthritis (CGA). When uric acid concentration in blood was above 7 mg/L or 420 $\mu\text{mol/L}$ (Hyperuricemia, HUA), monosodium urate (MSU) crystals might accumulate in joint capsule, bursa, cartilage, bone, or other periarticular tissues, which stimulated the synovial membrane of the joint, and produced pathological reactions such as synovial vasodilation, as well as increased permeability and leukocyte exudation (Acute gouty arthritis, AGA). However, some patients with HUA would not be transformed into AGA, only showing excessive serum uric acid[1], statistics show that up to 30% of AGA patients were provided with normal uric acid value[2,3]. Therefore, high uric acid concentration in blood was a necessary and insufficient condition for AGA. Clinically, AGA was

most frequently encountered in the major joints, especially in the first metatarsophalangeal joint, ankle, and foot joints. Other complications include chronic renal injury, ureteral calculi, and arthritis malformation [4,5]. Long-term intermittent repeated episodes of AGA would lead to the deposition of uratoma and eventually evolve into chronic gouty arthritis (CGA), and there would be a period of asymptomatic interlude (During the intermittenment period, DIP) between the evolution of AGA to CGA. The abridged general view of GA was provided in Figure 1.

Figure 1. Abridged general view of the evolution of gouty arthritis

According to the statistics, the incidence rate of GA were increasing year by year worldwide (about 1–2%, 2018), especially in developing countries [6,7]. With the deepening of research in recent years, GA has been gradually defined as an autoimmune inflammatory disease [8]. Some preliminary research conclusions on the pathogenesis of AGA and CGA have been made by researchers at present. The pathogenesis of AGA might be related to the activation of Toll-like receptors, NLRP3 inflammasomes or P2X7 receptors [9], while CGA might be related to the stimulation of the generation of extracellular neutrophil traps [10] or endoplasmic reticulum stress response [11]. Some scholars also analyzed the metabolites of gout, HUA patients and healthy volunteers by NMR [12] and IC-MS [13], and thus obtained the relevant biomarkers of gout and HUA.

The clinical diagnosis of GA was generally based on joint swelling[14], CT[15], smear test[16] and the description of patients, which was provided with uncertainties (MSU crystal smear of hyperuricemia was also presented positive results)[17]. At present, there are different treatment methods for GA patients at different stages, such as diet control, drug treatment and surgical treatment. However, in the absence of clear diagnostic markers, there is still a lack of methods to prevent the occurrence of GA and predict the evolution trend of GA. Therefore, white blood cell (WBC), C reactive protein (CRP), uric acid (UA), blood urea nitrogen (BUN), creatinine (Cre), hemoglobin (Hem), erythrocyte sedimentation rate (ESR), high/low-density lipoprotein (HDL/LDL), total cholesterol (TC), triglyceride (TG), demographic (age, body mass index(BMI), sex) and living habit(smoking and alcohol drinking habit) had been considered as the candidate for the risk factors of the progression of GA in this reaserch. For most of the patients, the symptoms of GA and Rheumatoid Arthritis(RA) are similar, so it is easy to ignore the condition and delay the optimal treatment time. Therefore, we also distinguished the two by the difference of biochemical indicators. Furthermore, principal component analysis (PCA), orthogonal partial least squares discrimination analysis (OPLS-DA), non-repetitive one-way ANOVA[18], correlation analysis[19] and multiple logistic regression analysis[20] were used to screen the important indicators affecting GA in each stage, and to distinguish GA and RA. Finally, multiple logistic regression was used to establish a clinical diagnosis model [21], so as to improve the success rate of clinical diagnosis and prediction. The overview of study design was indicated in Figure S1.

2. Results

2.1 Statistical analysis of patient information and serum biochemical indicators

Demographic data, living habits, comorbidities, disease durations and medical situations were collected through questionnaires and case history. UA , WBC, CRP, BUN, creatinine, hemoglobin, ESR, HDL/LDL, TC and TG of participators in each group were determined by the automatic biochemistry analyzer(ACA), the outcome of statistical analysis was provided in Table 1(training set) , and the trend of biochemical indicators in the five stages was indicated by Box-plot(Figure 2). The RA group (n=32) was added in this study in order to find a method that could be used to distinguish CGA from RA.

2.2 Principal component analysis and orthogonal partial least squares discrimination analysis

Multivariate analysis was performed on SIMCA-P 16.1 with the raw data provided in Table S1. In order to figure out the intrinsic differences among the 6 groups (control, HUA, AGA, DIP, CGA and RA), principal component analysis (PCA) and Orthogonal Projections to Latent Structures- Discriminant Analysis (OPLS-DA) were utilized to distinguish. The PCA score plots (Figure 3A) showed a satisfactory separating effect of data among the 6 groups, while the $R^2X=1.00$ and $Q^2=0.978$. Moreover, OPLS-DA score plots (Figure 3B) showed a better consequence, while the $R^2X=0.707$, $R^2Y=0.483$ and $Q^2=0.453$. As can be seen from Figure 3, the biochemical index profile of each group was significantly different from that of the control group, with the greatest difference between the CGA group and the control group. Furthermore, the serum biochemical profiles of CGA and RA were relatively different, indicating that the biochemical indicators selected in this study could adequately distinguish CGA from RA.

Table 1 Baseline characteristics (demographic, living habit, medication, blood biochemical values) of the HUA, AGA, DIP, CGA, RA and control groups in training set (n=379)

Parameter	Control	GA (n=267)				RA
	(n = 80)	HUA (n=62)	AGA (n=69)	DIP (n=74)	CGA (n=62)	(n=32)
Age (years)	52.1±9.3	46.5±15.8	49.6±17.6	44.2±16.2	69.0±6.8**#	56.3±10.9
BMI (kg/m ²)	22.5±2.6	26.1±1.0**#	28.9±1.9**#	25.1±1.4**#	21.3±1.0##	21.7±1.8
UA (μmol/L)	147.0±89.4##	481.9±60.8**#	473.0±120.7**#	420.0±104.8**#	338.7±76.6**#	284.3±93.4**
BUN (mmol/L)	5.0±1.0	5.0±1.4	5.8±3.7	5.1±1.7	8.7±1.9**#	5.8±2.3
Creatinine (μmol/L)	75.7±15.0#	82.0±16.3*#	101.9±16.7*	104.0±16.1*	211.3±31.0**#	104.2±21.8*
Hemoglobin (g/L)	134.8±8.8	132.8±7.0	135.6±8.5	133.8±7.8	133.2±7.7	135.4±8.5
ESR (mm/h)	14.8±3.0##	18.1±1.4#	30.9±8.0**#	16.7±1.9#	24.1±4.6*	24.9±5.3**
HDL (mmol/L)	1.5±0.6	1.6±0.5	1.1±0.2**#	1.0±0.2**#	1.3±0.3	1.5±0.4
LDL (mmol/L)	2.8±0.7	2.5±0.6	2.9±0.9	3.0±0.6	3.0±0.9	2.8±0.8
WBC (10 ⁹ /L)	6.0±1.5##	5.7±0.9#	8.4±2.8**	7.3±1.8*	7.7±1.2**	7.5±2.6**
TC (mmol/L)	5.2±0.9	4.9±1.0	4.9±0.9	5.2±1.2#	5.4±0.8	4.7±1.1
TG (mmol/L)	1.3±0.7	2.6±1.2**#	1.9±0.9**#	2.4±1.5**#	2.1±1.1**#	1.2±0.4
CRP (Mg/L)	0.7±0.8##	0.9±0.4#	30.5±26.9**	6.8±17.9**	11.4±6.9**#	21.9±22.4**
RF	<20##	<20##	<20##	<20##	<20##	290.1±302.1**
Smoker (%)	34.4	37.5	55.1**#	57.3**#	73.3**#	20.8
Alcohol drinking habit (%)	25.0##	31.9##	40.6**#	42.7**#	70.0**#	4.2
Sex (female/male)	35/45	16/46	5/64	4/70	2/60	26/6
Disease duration (months)	/	5.5±3.5	2.0±1.5	1.5±1.0	> 132 (11 years)	> 60 (5 years)
ULT (febuxostat/benzbromarone/non)	(0/0/160)	(9/1/22)	(42/5/22)	(0/0/74)	(32/0/0)	(0/0/32)
Major Disease (tumour/CVD/non)	(0/0/160)	(1/3/28)	(0/0/69)	(0/0/74)	(3/24/3)	(1/2/29)

Data were presented as mean ± SD except where stated otherwise.

HUA: hyperuricemia; AGA: acute gouty arthritis; DIP: during the interictal period; CGA: chronic gouty arthritis; RA: rheumatoid arthritis; BMI: body mass index; WBC: white blood cell; CRP: C-reactive protein; RF: rheumatoid factors; UTL: uric acid lowering therapy; CVD: cardiovascular disease.

* $p < 0.05$, ** $p < 0.01$, compared to control; # $p < 0.05$, ## $p < 0.01$, compared to the RA group.

The calculation results of p value were shown in Table S3.

2.3 One-way ANOVA

The non-repetitive one-way ANOVA results of serum samples from different groups of patients were shown in Table S3, and the variance summary of biochemical indicators in each group was provided in Table 2.

2.4 Correlation analysis of factors influencing the course of GA

The corrplot package written by our research group in R language was applied for correlation analysis of biochemical indicators associating with progression of GA. The matrix heat map of correlation were shown in Figure 4, and detailed analysis data were provided in Table S4. Here, we assume that the GA process is gradually evolved from the control group, HUA group, AGA group, DIP group and CGA group from light to heavy. As can be seen from Figure 4, LDL and TC were significantly correlated, indicating that the correlation was not affected by the disease and had little significance for the evolution of GA. In addition, UA, creatinine, WBC, CRP, BUN, ESR, BMI and

HDL were significantly correlated with the progression of GA, among which UA, creatinine, WBC, CRP, BUN, and ESR were remarkably positively correlated with the progression of GA, while HDL and BMI were remarkably negatively correlated with it.

Figure 2. The eight selected blood biochemical indexes for the identification of the five GA stages and RA. A–L Boxplots showed the blood concentration of the eight indexes (left y-axis) including WBC, CRP, uric acid, urea nitrogen, HDL/LDL, cholesterol, triglyceride, creatinine, hemoglobin, BMI and ESR across all five GA stages and RA. In the boxplot, the median was expressed by the center line, 75% was represented by the upper bound of the box, 25% was expressed by the lower bound of the box, minimum is expressed by the lower whisker, and maximum was expressed by the upper whisker. Dot plots and lines showed the trend of indexes across all five GA stages and RA.

Figure 3. PCA and OPLS-DA results of serum biochemical profile of 6 groups. A was PCA score plots, B was OPLS-DA score plots, C was variable important in projection (VIP) value of each biochemical index in OPLS-DA model, D was 3D OPLS-DA score plots. The confidence interval is 95%. Creatinine, Age, ESR, BMI, Hemoglobin, WBC, CRP and UA contributed significantly to the difference between groups (VIP > 1).

Table 2 Variance summary of biochemical indicators in each group

Groups \ Indexs	WBC	CRP	BUN	UA	HDL	LDL	TC	TG	BMI	Cre	Hem	ESR
Control	2.719	0.685	0.871	4126.055	0.455	0.395	0.610	0.350	6.199	213.424	71.815	7.677
HUA	0.854	0.145	1.470	11787.984	0.211	0.467	0.774	1.699	1.271	300.712	58.759	2.432
AGA	6.536	459.313	7.364	13350.600	0.055	0.895	1.134	0.675	3.124	358.608	81.007	61.942
DIP	3.689	689.154	2.067	12651.144	0.033	0.383	1.213	2.077	2.024	236.446	62.674	2.674
CGA	5.697	129.271	15.506	6686.177	0.208	1.078	1.059	1.571	1.434	1009.090	71.542	25.492
RA	7.184	671.109	6.682	7324.392	0.153	0.715	1.239	0.163	3.842	646.156	71.917	27.575

Figure 4. The matrix heat map of correlation of biochemical indicators associating with progression of GA. Each dot represented a correlation between two indicators, and the dot size was proportional to the P value, the color gradation of the dots represents the magnitude of the correlation. The number at the bottom left represented the corresponding correlation coefficient with the dot. Blue indicated positive correlation, red indicated negative correlation. As can be seen, creatinine (0.62), ESR (0.36), BUN(0.32), WBC (0.31) and UA(0.23) had the strongest correlation with the progression of GA. Matrix heat maps of correlation in other forms generated by corplot package written by our research group in R language were shown in Figure S2.

2.5 Clinical diagnostic model associating with progression of GA based on serum biochemical indicators

MASS, PROC, ggplot2, mlogit, RMS and survival packages in R language were used to carry out ordinal multiple logistic regression for clinical diagnostic model associating with progression of GA. The independent variables were without influence upon regression coefficient and segmentation point of the dependent variables, and there was no multicollinearity between the independent variables. The dependent variable was transformed into the corresponding dummy variable before modeling, and then the continuous variables (Among the independent variables, “Sex” was the classified variable, and the rest were continuous variables) were standardized, followed by factor analysis. The training set data (n=379) was adopted to build the model, the five stages of GA(Ccontrol、HUA、AGA、DIP、CGA) were regarded as dependent variable, then logistic stepwise regression was performed for each stage. Meanwhile, combining the results of “2.1”-“2.4”, risk factors were screened to form the best logistic regression equation for each stage. The prediction probability of the sample under each regression equation was calculated, and the stage of GA in which the maximum probability was located was considered as the diagnosis. The verification set data (n=200) was adopted to evaluate and verify the model. Details of clinical diagnostic model of GA were provided in Table 3.

Table3 Details of clinical diagnostic model of GA

Model	Factors	OR	95%CI	P value	Regression Equation
Model Control	UA	0.97	0.96-0.98	2.21×10 ⁻¹⁵	$P_1 = \frac{1}{1 + e^{-Z_1}}$
Model HUA	CRP	8.28	3.91-19.34	1.93×10 ⁻⁷	$P_2 = \frac{1}{1 + e^{-Z_2}}$
	HDL	1.01	1.01-1.02	1.24×10 ⁻¹²	
	UA	0.95	0.92-0.97	6.29×10 ⁻⁷	
Model AGA	WBC	1.05	1.03-1.07	1.85×10 ⁻⁶	$P_3 = \frac{1}{1 + e^{-Z_3}}$
	UA	1.01	1.01-1.02	4.17×10 ⁻⁸	
	ESR	1.33	1.24-1.45	3.20×10 ⁻¹³	
Model DIP	WBC	1.20	1.01-1.43	3.78×10 ⁻²	$P_4 = \frac{1}{1 + e^{-Z_4}}$
	HDL	0.03	0.007-0.09	5.33×10 ⁻⁸	
	UA	1.01	1.003-1.009	1.15×10 ⁻⁶	
Model CGA	ESR	0.73	0.66-0.80	2.76×10 ⁻⁹	$P_5 = \frac{1}{1 + e^{-Z_5}}$
	BUN	1.83	1.58-2.15	2.52×10 ⁻¹⁴	
	ESR	1.04	1.00-1.09	5.29×10 ⁻²	

OR= odds ratio, indicates the strength of the association between factors and models;

CI= confidence interval; P in regression equations: prediction probability.

Z1-Z5 in regression equations: Z1=5.38-0.0251×UA, Z2=4.12+2.11×HDL+0.0114×UA-0.0560×CRP, Z3=14.8+0.0464×CRP+0.0132×UA+0.287×ESR, Z4=4.92+0.184×WBC-3.52×HDL+0.00608×UA-0.310×ESR, Z5=6.58+0.602 ×BUN+0.0436×ESR

2.6 Evaluation and verification of clinical diagnostic model associating with progression of GA

Then, the validation set data was substituted into the above models, and the predicted results were outputted and compared with the actual result (Table S2). The receiver operating characteristic (ROC) curves of five models were shown in Figure 5. Total area under the curve (AUC) of clinical diagnostic model associating with progression of GA was 0.9534, and Kappa coefficient was applied to evaluate the consistency between the predicted results of the model and the actual results. Encouragingly, Kappa coefficient of the clinical diagnostic model was 0.80, which indicated a substantial consistency of this model (The magnitude of Kappa coefficient could be divided into five degrees, 0.0~0.20: slight consistency, 0.21~0.40: fair consistency, 0.41~0.60: moderate consistency, 0.61~0.80: substantial consistency, 0.81~1: almost perfect consistency). It shows that the model has high accuracy and reliability.

Figure 5. The receiver operating characteristic (ROC) curves of five models. The area under the curve (AUC) of five models were 0.9814 (Control), 0.9288(HUA), 0.9752(AGA), 0.9056(DIP), 0.9759(CGA).

2.7 Visualizing of clinical diagnostic model associating with progression of GA

In order to get more application of GA diagnostic model in clinical practice,, the nomogram of the clinical diagnostic model associating with progression of GA was made by adopting the r language rms package (Figure 6). According to the regression coefficient of each influencing factor in the multiple regression model, each influencing factor is scored, and the summation function of the score is converted into the probability of the occurrence of the outcome event, and the stage of the maximum probability is the diagnosis result. This method can transform the complex regression equation into a simple and visual chart, making the results of the prediction model more readable and useful.

Figure 6. The nomogram of five models. A: Control group, B: HUA group, C: AGA group, D: DIP group, E: CGA group.

3. Discussion

Statistical studies showed that the prevalence of GA increased year by year, especially in developing countries, with a 2% increase in the prevalence of GA from 2017 to 2019[22-23]. The incidence of GA in European and American countries was 0.13% ~ 0.37%, and the annual incidence is 0.20% ~ 0.35%[24]; the prevalence of HUA in the general population in China was about 10%[25]. So far, more than 80% of the published articles on disease prediction models for arthritis were associated with RA[26-29] according to our literature review and document retrieval. Thus it could be seen that GA was not a research hotspot in the field of arthritis, which was not commensurate with its high incidence. In order to reverse this abnormal phenomenon, the study of clinical diagnostic model and IRF associating with progression of GA based on Serum biochemical indicators was implemented in this article.

We found a strong correlation between LDL and TC in all groups(LDL/TC in control, HUA, AGA, DIP and CGA group: 0.54, 0.51, 0.59, 0.58 and 0.55), which was consistent with literature reports[30]. The correlation was not affected by GA progression, which indicated that LDL and TC were not IRF of progression of GA. Moreover, according to line charts in Figure 2, hemoglobin tended to flatten out in the five stages of GA. Therefore, LDL, TC and hemoglobin were excluded in the establishment of diagnostic model.

Boxplots C and I of Figure 2 clearly indicated that BUN and creatinine showed an significant elevation in CGA group compared with the other five groups. The kidney was the main organ for excreting BUN, as with serum creatinine, BUN could be in the normal range in the early stages of renal function impairment. However, BUN and creatinine would rise rapidly when the glomerular filtration rate(GFR) drops below 50% of normal. Studies have shown that chronically high levels of uric acid in the blood could significantly increase the risk of kidney disease[31], and 5.6% of 13338 participants(mean serum uric acid = 5.9 ± 1.5 mg/dL) had incident kidney disease defined by GFR decrease of more than 30% over 8.5 years[32]. Another population-based cohort study showed that patients with CGA who were treated with urate-lowering therapy (ULT) had a greater risk of incident chronic kidney disease (CKD)[33]. Therefore, BUN and creatinine could be regarded as IRFs for the progression of GA, and the significant rising of BUN and creatinine indicates that the course of GA has entered the CGA stage. However, there was no significant difference in BUN and creatinine between the control, HUA, AGA and DIP group, so the establishment of a clinical diagnostic model for the progression of GA requires the contribution of other indicators. Besides, BUN and creatinine could be considered as a the diagnostic indicators of CGA and was a serum biochemical indicator that could distinguish CGA from RA in addition to rheumatoid factors(RF).

Moreover, as can be seen in Figure 3, the biochemical profile of each group was significantly different from that of the control group, with the greatest difference between the CGA group and the control group, as well as creatinine, age, ESR, BMI, hemoglobin, WBC, CRP and UA contributed significantly to the difference between groups ($VIP > 1$). Furthermore, the serum biochemical profiles of CGA and RA were relatively different, indicating that the biochemical indicators selected in this study could adequately distinguish 6 groups including 5 GA stages and RA. On this account, we suppose that the bias of the overall profile of the patient's serum biochemical indicators to a certain stage may lead to the appearance of symptoms. Therefore, we believe that the previous single indicator diagnosis method should be improved, and turn to adopt multiple indicators for clinical comprehensive diagnosis and prediction. Meanwhile, according to the results of correlation analysis, UA, creatinine, WBC, CRP, BUN, TG, ESR and HDL were significantly correlated with the progression of GA, among which creatinine (0.62), ESR (0.36), BUN(0.32), WBC (0.31) and UA(0.23) had the strongest correlation with it, this was basically consistent with the results of multivariate statistics.

The training set data (n=379) was adopted to build the clinical diagnostic model associating with progression of GA (consists of five phases), and the verification set data (n=200) was adopted to evaluate and verify it. AUCs of five models were 0.9814 (Control), 0.9288(HUA), 0.9752(AGA), 0.9056(DIP), 0.9759(CGA). The Kappa coefficient applied to evaluate the consistency was 0.80, indicated a substantial consistency of this model. Furthermore, The visualization of the model is realized by using nomograms, which facilitates more clinical application of the model. Both doctors and patients can judge the development and changes of GA according to the model, which provides an effective reference for distinguishing the development stages of gouty arthritis.

After this study, non-targeted metabolomics analysis will be performed on the serum samples of GA patients collected above to obtain potential diagnostic biomarkers related to the progression of GA, including absolutely qualitative and quantitative verification. Then, the changes of biomarkers and serum biochemical indexes in different progressions of GA were found to optimize the clinical diagnostic model of GA based on serum biochemical indexes proposed in this paper. It is expected

to develop a new clinical diagnosis method of GA similar to the "GA diagnostic kit", which can help clinicians predict the progress of GA faster and more accurately.

4. Methods and ethical approval

4.1 Study design and measurement

This prospective study was implemented during 30 November 2017 and 10 November 2019 at Zhejiang Provincial People's Hospital(ZPPH), the largest comprehensive first-class hospital in Zhejiang province of China. All participants were categorized as healthy control, HUA, AGA, DIP and CGA based on the level of serum uric acid and the definition of GA as follows (based on 2015 ACR-EULAR Gout Classification Criteria)[34]: (1) healthy control was defined as serum uric acid of 150 ~ 416 $\mu\text{mol/L}$ in males or 80 ~ 357 $\mu\text{mol/L}$ in female and never had GA or other types of arthritis, such as rheumatic arthritis, rheumatoid arthritis, infectious arthritis, osteoarthritis and psoriatic arthritis. (2) HUA was defined as serum uric acid of $> 416 \mu\text{mol/L}$ in males or $> 357 \mu\text{mol/L}$ in female. (3) AGA was defined as persistent swelling and intense pain in the peripheral joints or bursae, meanwhile, the symptoms peaked within 24 hours and resolved within 14 days. (4) DIP was defined as having been over 4 weeks since the last AGA attack and had not received uric acid lowering therapy (ULT). (5) CGA was defined as joint swelling, pressing pain, deformity, dysfunction and hypodermic tophus. (6) all the patients' synovial fluid in HUA and AGA groups were identified by polarized-light microscopy to ensure the presence of MSU crystals.

In addition to hyperuricemia, the onset of GA was usually associated with age, BMI, sex, smoking and alcohol drinking habit[35]. Therefore, demographic data, living habits, comorbidities (tumor and cardiovascular disease), disease durations and medical situations were collected through questionnaires and case history, which was provided in Table 1. Serum uric acid levels and other blood biochemical indexes of participants in each group were determined by the CHEMIX-180 automatic biochemistry analyzer (Sysmex Corp., Kobe, Japan). Each subject was visited and collected serum only once during the study period.

4.2 Study population and Statistical analysis

A total of 579 serum specimens were collected from ZPPH, including 379 in training set: 80 healthy volunteers (21.11%), 62 patients with HUA (16.36%), 69 patients with AGA (18.21%), 74 patients with DIP (19.535%), 62 patients with CGA (16.36%), and 32 patients with RA (8.43%). 200 in validation set: 80 healthy volunteers (40.00%), 30 patients with HUA (15.00%), 30 patients with AGA (15.00%), 30 patients with DIP (15.00%), and 30 patients with CGA (15.00%). About 70% of the patients were inpatients and another 30% were outpatients, the acute episodes of AGA lasted no less than 3 days and the course of CGA was more than 11 years. Among the 207 patients with GA (mean value of age = 49.23 ± 17.90), 188 were male (90.8%) and 19 (9.2%) were female. GA patients had a higher average BMI (25.88 ± 3.01 vs. $22.9 \pm 2.7 \text{ kg/m}^2$; $p = 0.01$) and uric acid (420.81 ± 116.86 vs. 158.2 ± 89.0 ; $p = 0.001$) than those in the control group. Meanwhile, the percentage of GA patients with smoking and alcohol drinking habits was much higher than the healthy volunteers (smoker: 34.4% vs. 55.8%; drinker: 25.0% vs. 46.3%). Statistical description of other biochemical indicators in all serum specimens was shown in table 1, and the trend of biochemical indicators in the five disease progressions was indicated by Box-plot (Figure 2). In

addition, the raw data of biochemical indicators in all serum specimens was provided in Table S1 (training set) and S2(validation set). The difference in measurement with a two-sided alpha level of 0.05 could be ensured by this sample size using Fisher's exact test. Simca.16.1(Umetrics inc., Sweden) was used for principal component analysis and partial least square analysis in 6 groups (control group, HUA group, AGA group, DIP group, CGA group). Using SPSS23.0 (SPSS Inc, Chicago, Illinois, USA) for different group of patients serum samples of repetitive single factor analysis of variance. Afterwards Pearson correlation coefficients (PCCs) was conducted on the indexes that may affect the progression of GA by using R language packages(corrplot). By applying the code written by our research team and combining 7 R language packages (quality plan, pROC package, ggplot2 package, mlogit package, RMS package, corrplot package and survival package), we carried out multiple logistic regression analysis, established GA diagnosis model, and conducted verification and visualization.

5. Conclusion

In this study, serum biochemical indicators from clinical assessment were adopted to establish clinical diagnostic model and serum biochemical profile associating with progression of GA. It is the first study to investigate into the serum biochemical profile at different stages of gouty arthritis. An obvious difference in serum biochemical profile among stages of GA was found in this paper, which could effectively distinguish them. We suppose that the bias of the overall profile of the patient's serum biochemical indicators to a certain stage may lead to the appearance of symptoms. A simple evaluation tool was developed, we have reason to believe it will be useful for clinicians and researchers wishing consolidate their clinical diagnosis in regard to the stage and tendency of GA patients. This clinical diagnostic model could be applied clinically and in research to improve accuracy of identification and prediction of these stages of GA patients.

Abbreviations

GA: Gouty arthritis; RA: Rheumatoid arthritis; HUA: Hyperuricemia; AGA: Acute gouty arthritis; DIP: During the intermittent period; CGA: Chronic gouty arthritis; AUC: Area under curve; BMI: Body mass index; CRP: C reactive protein; CV: Coefficient of variation; ANOVA: Analysis of variance; CVD: Cardiovascular disease; MSU: Monosodium urate; MVA: Multivariate analysis; OPLS-DA: Orthogonal projections to latent structures-discriminant analysis; PCA: Principal component analysis; RF: Rheumatoid Factors; ROC: Receiver operating characteristic curve; ULT: Urate-lowering therapy; WBC: White blood cell. JUTCM: Jiangxi University of Traditional Chinese Medicine; ZPPH: Zhejiang Provincial People's Hospital.

Funding:

This work was supported by grants (No. 81560636, 81760702) from Natural Science Foundation of China, National Key R&D Program of China (No.2019YFC1712300), Jiangxi Province 5511 innovative talent Project [20165BCB19009], Nan-chang innovative talent team (No. [2016]173) and Project of Jiangxi Provincial Education Department (YC2018-B073).

Availability of data and materials

The following data are available in “**Supplementary Materials**”, Table S1: Raw data of biochemical indicators in serum specimens of training set (n=379) , Table S2: Raw data of biochemical indicators in serum specimens of validation set (n=200) , Table S3: The result of one-way ANOVA, Table S4: Detailed correlation analysis data,

Figure S1: The overview of study design, Figure S2: Matrix heat maps of correlation in 4 forms generated by corrrplot package written by our research group in R language.

The data that support the findings of this study are available from JUTCM but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available from the authors upon reasonable request and with permission of JUTCM.

Authors' contributions:

Conceptualization, Shang Lyu, Shilin Yang and Yulin Feng; Data curation, Shang Lyu; Formal analysis, Ruowen Ding; Funding acquisition, Yulin Feng; Investigation, Wanyuan chen, Peng Liu and Lijuan Wang; Project administration, Yulin Feng; Software, Ruowen Ding; Supervision, Yi Rao; Validation, Hui OuYang; Visualization, Ruowen Ding; Writing – original draft, Shang Lyu; Writing – review & editing, Shang Lyu.

Ethics approval and consent to participate

The medical ethics committee of ZPPH approved this study (acceptance number: KY2018048, approval number: 2018KY046), the ethics committee was constituted and operated in accordance with the principles of SFDA/GCP and the Declaration of Helsinki. Health screening of participants in the study were enrolled at ZPPH from 2018 to 2019. Demographic data, living habits, comorbidities, disease durations and medical situations were recorded in a normative questionnaire.

Consent for publication

The publication of all case reports has been consent, and the personal data contained in the manuscript have been consent to publish by the patient.

Competing interests

The authors declare that they have no competing interests.

References

1. Merriman, T.R.; Choi, H. K.; Dalbeth, N. The genetic basis of gout. *Rheum Dis Clin North Am*, 2014, 40, 279-290.
2. Schumacher HR (1996) Crystal-induced arthritis: an overview. *Am J Med*, 100(2a), 46–52.
3. Chayawee M, Thitirat B(2017) An investigation of the independent risk factors that differentiate gout from pseudogout in patients with crystal-induced acute arthritis: a cross-sectional study. *Rheumatol Int* (06), doi.org/10.1007/s00296-017-3891-2
4. Dalbeth, N.; Merriman, T.R.; Stamp, L.K. Gout. *Lancet* 2016, 388, 2039–2052.
5. Roddy, E.; Menon, A.; Hal, A.; Datta, P.; Packham, J. Polyarticular sonographic assessment of gout: A hospital-based cross-sectional study. *Joint Bone Spine* 2013, 80, 295–300.
6. Kuo, C.F.; Grainge, M.J.; Zhang, W.; Doherty, M. Global epidemiology of gout: Prevalence, incidence and risk factors. *Nat. Rev. Rheumatol.* 2015, 11, 649–662.
7. Neogi, T. Gout. *N Engl. J. Med.* 2011, 364, 443–452.
8. Richette, P.; Bardin, T. Gout. *Lancet* 2010, 375, 318–328.
9. Flávio, A. A.; Costa, V. V.; Tavares, L. D. NLRP3 inflammasome-mediated neutrophil recruitment and hypernociception depend on leukotriene B4 in a murine model of gout. *Arthritis & Rheumatism*, 2012, 64(2),474-484.

10. Schauer, C.; Anko, C. Aggregated neutrophil extracellular traps limit inflammation by degrading cytokines and chemokines. *Nature Medicine*, 2014, 20(5), 511-517.
11. Tanaka, K.; Kaji, H.; Yamaguchi, T. Involvement of the osteoinductive factors, Tmem119 and BMP-2, and the ER stress response PERK-eIF2 α -ATF4 pathway in the commitment of myoblastic into osteoblastic cells. *Calcif Tissue Int*, 2014, 94(4), 454-464.
12. Zhang, Y. N.; Zhang, H. Z.; Chang, D.; Guo, F. C.; Pan, H. Z.; Yang, Y. X. Metabolomics approach by ¹H NMR spectroscopy of serum reveals progression axes for asymptomatic hyperuricemia and gout. *Arthritis Research & Therapy*, 2018, 20(111).
13. Cui, L.; Liu, J.; Yan, X. M. Identification of Metabolite Biomarkers for Gout Using Capillary Ion Chromatography with Mass Spectrometry. *Analytical chemistry*, 2017, 89(21), 11737-11743.
14. Muangchan, C.; Bhurihirun, T. An investigation of the independent risk factors that differentiate gout from pseudogout in patients with crystal-induced acute arthritis: A cross-sectional study. *Rheumatol. Int.* **2017**, 38, 89-95.
15. Finkenstaedt, T.; Manoliou, A.; Toniolo, M.; Higashigaito, K.; Andreisek, G.; Guggenberger, R.; Michel, B.; Alkadhi, H. Gouty arthritis: The diagnostic and therapeutic impact of dual-energy CT. *Eur. Radiol.* **2016**, 26, 3989-3999.
16. Robier, C.; Stettin, M.; Quehenberger, F.; Neubauer, M. Cytospin preparations are superior to common smears in the detection of monosodium urate crystals in low-cellular synovial fluids. *Clinical Rheumatology*, 2014, 33(12), 1797-1800.
17. Lyu, S.; Ding, R. W.; Liu, P.; OuYang, H.; Feng, Y. L.; Rao, Y.; Yang, S. L. LC-MS Analysis of Serum for the Metabolomic Investigation of the Effects of Pulchrenoside b4 Administration in Monosodium Urate Crystal-Induced Gouty Arthritis Rat Model. *Molecules*, 2019, 24(3161), doi:10.3390.
18. Caroline, M. F.; Gillian, M. M.; Nicholas, S.; Trevor, W. S. Inflammatory status and kynurenine metabolism in rheumatoid arthritis treated with melatonin. *British Journal of Clinical Pharmacology*. 2007, 64(4), 517-526.
19. Rezaei, H.; Klint, E.; Hammer, H.B.; Terslev, L.; Agostino, M. A.; Kisten, Y.; Arnaud, L. Analysis of correlation and causes for discrepancy between quantitative and semi-quantitative Doppler scores in synovitis in rheumatoid arthritis. *Rheumatology*, 2017, 56(2), 255-262.
20. Bécède, M.; Alasti, F.; Gessl, I.; Haupt, L.; Kerschbaumer, A.; Landesmann, U. Risk profiling for a refractory course of rheumatoid arthritis. *Seminars in Arthritis and Rheumatism*, 2019, 49(2), 211-217.
21. Wang, Y. H.; Zhu, R.; Xiao, J.; Davis, J. C.; Mandema, J. W.; Jin, J. Y. Short-Term Efficacy Reliably Predicts Long-Term Clinical Benefit in Rheumatoid Arthritis Clinical Trials as Demonstrated by Model-Based Meta-Analysis. *J CLIN PHARMACOL*, 2016, 56(7), 835-844.
22. Zobbe, K.; Prieto-A, D.; Cordtz, R.; Højgaard, P.; Hindrup, J.; Kristensen, L. Secular trends in the incidence and prevalence of gout in Denmark from 1995 to 2015: a nationwide register-based study. *Rheumatology*, 2019, 58(5), 836-839.
23. Wändell, P.; Andreasson, A.; Hagström, H.; Kapetanovic, M. C.; Carlsson, A. C.; Kapetanovic, M. C. The use of anthropometric measures in the prediction of incident gout: results from a Swedish community-based cohort study. *SCAND J RHEUMATOL*, 2019, 48(4), 294-299.
24. Fredrik, N.; Laura, H.; Robert, M.; Javier, N.; Chris, S.; Lalitha, A. Comorbidity Burden in Trial-Aligned Patients with Established Gout in Germany, UK, US, and France: a Retrospective Analysis. *Advances in therapy*, 2016, 33(7), 1180-1198.
25. Chen, Y. Z.; Tang, Z. Z.; Huang, Z. Y.; Zhou, W. W.; Li, Z. Y.; Li, X. P. The prevalence of gout in mainland China from 2000 to 2016: a systematic review and meta-analysis. *Journal of Public Health*, 2017, 25(7), 521-529.

26. Ian, C. S.; Sophia, S.; Rachael, T.; Paola, F.; Ann, W. M.; Anne, H. Prediction model for rheumatoid arthritis: modelling 46 genetic risk variants with smoking. *Lancet*, 2013, *381*, 93.
27. Arana-Guajardo, A.; Perez-Barbosa, L.; Vega-Morales, D. ;Riega-Torres, J.; Esquivel-Valerio, J.; Garza-Elizondo, M. Application of a prediction model for the progression of rheumatoid arthritis in patients with undifferentiated arthritis. *Reumatol Clin*, 2014, *10*(6).
28. Tamarah, D.; Sellam, J.; Agca, R.; Vosslander, S.; Witte, B.; Tsang-A-Sjoe, M. A multi-parameter response prediction model for rituximab in rheumatoid arthritis. *Joint Bone Spine*, 2018, *58*(2), 219-226.
29. Nathan, V.; Stephen, X.; Daniel, A.; William, C.; Josef, S. A pilot risk model for the prediction of rapid radiographic progression in rheumatoid arthritis. *Rheumatology*, 2009, *48*(9), 1114-1121.
30. Kamel, S.; Richards, B.; Dufresne, L.; Engert, J.; Thanassoulis, G. Total cholesterol, low-density lipoprotein cholesterol, or high density lipoprotein and bone health: systematic review and mendelian randomization. *Can J Cardiol*, 2015, *31*(10), 304.
31. Obermayr, P.; Temml, C.; Gutjahr, G.; Knechtelsdorfer, M.; Oberbauer, R.; Klauser, R. Elevated uric acid increases the risk for kidney disease. *J Am Soc Nephrol*, 2008, *19*(12), 2407-2413.
32. Weiner, D.E.; Tighiouart, H.; Elsayed, E. F.; Griffith, J. L.; Salem, D. N.; Levey, A. S.; Uric acid and incident kidney disease in the community. *J Am Soc Nephrol*, 2008, *19*(6), 1204-1211 .
33. Roughley, M.; Sultan, A. A.; Clarson, L.; Muller, S.; Whittle, R.; Belcher, J , *et al.* Risk of chronic kidney disease in patients with gout and the impact of urate lowering therapy: a population-based cohort study. *Arthritis Rese Ther*, 2018, *20*(1), 243.
34. Neogi, T; Jansen, T; Dalbeth, N; Fransen, J; Schumacher,H; Berendsen, D. et al. 2015 Gout Classification Criteria: An American College of Rheumatology/European League Against Rheumatism Collaborative Initiative. *Arthritis & rheumatology*, 2015, *67*(10), 2557-2568.
35. Chayawee, M; Thitirat, B. An investigation of the independent risk factors that differentiate gout from pseudogout in patients with crystal-induced acute arthritis:a cross-sectional study. *Rheumatology International*, 2018, *38*(1), 89-95.

Figures

Figure 1

Abridged general view of the evolution of gouty arthritis

Figure 2

The eight selected blood biochemical indexes for the identification of the five GA stages and RA. A–L Boxplots showed the blood concentration of the eight indexes (left y-axis) including WBC, CRP, uric acid, urea nitrogen, HDL/LDL, cholesterol, triglyceride, creatinine, hemoglobin, BMI and ESR across all five GA stages and RA. In the boxplot, the median was expressed by the center line, 75% was represented by the upper bound of the box, 25% was expressed by the lower bound of the box, minimum is expressed by the lower whisker, and maximum was expressed by the upper whisker. Dot plots and lines showed the trend of indexes across all five GA stages and RA.

Figure 3

PCA and OPLS-DA results of serum biochemical profile of 6 groups. A was PCA score plots, B was OPLS-DA score plots, C was variable important in projection (VIP) value of each biochemical index in OPLS-DA model, D was 3D OPLS-DA score plots. The confidence interval is 95%. Creatinine, Age, ESR, BMI, Hemoglobin, WBC, CRP and UA contributed significantly to the difference between groups (VIP $\geq$ 1).

Figure 4

The matrix heat map of correlation of biochemical indicators associating with progression of GA. Each dot represented a correlation between two indicators, and the dot size was proportional to the P value, the color gradation of the dots represents the magnitude of the correlation. The number at the bottom left represented the corresponding correlation coefficient with the dot. Blue indicated positive correlation, red indicated negative correlation. As can be seen, creatinine (0.62), ESR (0.36), BUN(0.32), WBC (0.31) and UA(0.23) had the strongest correlation with the progression of GA. Matrix heat maps of correlation in other forms generated by corrplot package written by our research group in R language were shown in Figure S2.

Figure 5

The receiver operating characteristic (ROC) curves of five models. The area under the curve (AUC) of five models were 0.9814 (Control), 0.9288(HUA), 0.9752(AGA), 0.9056(DIP), 0.9759(CGA).

Figure 6

The nomogram of five models. A: Control group, B: HUA group, C: AGA group, D: DIP group, E: CGA group.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryMaterial.pdf](#)