Validating traditional Chinese syndrome features in varied stages of chronic gastritis malignant transformation: study protocol for a cross-sectional study

Yin Zhang,1,2,3 Yue Liu,1 Rui Song,4,5 Li Zhang,6 Zeqi Su,1 Yannan Li,7 Runhua Chen,1,8 Ning Shi,1,9 Xia Zhao,10 Shiyu Du,11 Xia Ding1

ABSTRACT

Introduction The transition from chronic non-atrophic gastritis (CNAG) to chronic atrophic gastritis (CAG) and gastric carcinoma (GC) is regarded as a representative disease model of gastric mucosa malignant transformation led by uncontrolled inflammation. Traditional Chinese medicine (TCM) syndrome-targeted therapies have been applied in treating chronic gastritis (CG) malignant transformation in China with satisfying efficacy. This study aims to validate TCM syndrome features in each stage of CG malignant transformation. The findings may shed light on the TCM hypothesis of CG malignant transformation, and thus optimise syndrome-targeted treatment strategies of CNAG, CAG and GC, respectively.

Methods and analysis The present study is a cross-sectional study conducted in China. A total of 2000 eligible patients, including 500 CNAG cases, 1000 CAG cases and 500 GC cases, will be recruited from four TCM hospitals. Primary outcome measures include the prevalence of TCM syndrome patterns in varied stages of CG malignant transformation. Secondary outcome measures include prevalence and severity of all the presenting signs and symptoms collected by using TCM four diagnostic methods. Descriptive analysis, comparative analysis and correlation analysis of all the measurement data will be performed by biostatisticians. Unsupervised data mining analyses, including exploratory factor analysis, association rule analysis, hierarchical clustering analysis, complex system entropy clustering analysis, and so on, will also be performed by data scientists respectively for in-depth analyses of TCM syndrome-related indicators.

Ethics and dissemination The protocol has been approved by the Ethical Review Board of Dongzhimen Hospital Affiliated to Beijing University of Chinese Medicine (No ECPJ-BDY-2014-02). All the study outcomes will be disseminated through national conference reports and in the meantime published on peer-reviewed journals.

Trial registration number NCT03314038; Pre-results.

STRENGTHS AND LIMITATIONS OF THIS STUDY

To the authors’ knowledge, this is among the first reported large population-based studies on validating traditional Chinese medicine (TCM) syndrome features in varied stages of chronic gastritis malignant transformation.

A large sample size of 2000 participants can be achieved. Full-scale TCM and modern medicine indicators can be accurately collected.

Both statistical analysis and unsupervised data mining strategies will be applied by research personnel parallel to reinforce complementary strengths.

This is not a prospective study, and thus longitudinal data could not be collected by research personnel to speculate causal relationship.

Selection bias and information bias may exist.

INTRODUCTION

Chronic gastritis (CG), persistent inflammation of the gastric mucosa with no specific clinical manifestations, has been regarded as the most common gastrointestinal disease with an incidence of 70% among the adult population in China.1 Chronicity and recurrence of chronic non-atrophic gastritis (CNAG) may lead to chronic atrophic gastritis (CAG), which is characterised by gastric mucosal atrophy and usually accompanied by metaplasia or dysplasia.2 As the foremost step of CG malignant transformation, CAG is acknowledged as premalignant lesion of gastric carcinoma (GC). The sequential stages of inflammation, metaplasia, dysplasia and carcinoma have demonstrated a well-established tumorigenesis model of intestinal type of GC led by uncontrolled inflammation.2–4 Statistics have shown that there are 951600 annual new cases of GC worldwide.5 China has been reported as one of the countries with largest population of both CG and GC sufferers.6 Considering the severe burden of above-mentioned diseases, active
intervention on CG malignant transformation process is highly needed.6-8

Traditional Chinese medicine (TCM), as one of the most important and time-honoured complementary medicine approaches worldwide, has developed unique theories of aetiology and diagnosis since ancient time. According to TCM theory, syndrome is defined as a categorised pattern of symptoms and signs in a patient at a specific stage during the course of specific diseases, and is considered the most important unit for evaluating pathogenesis. TCM practitioners have been applying the central principles of syndrome pattern differentiation in clinical practice to identify the physical condition of patients and the pathogenesis of diseases with a history of more than 2500 years. Nowadays, TCM physicians keep on using the prestigious and traditional ‘four diagnostic methods’ of looking, listening/smelling, asking and palpating to collect the information of symptoms and signs on patients comprehensively. Full-scale assessment of aforementioned clinical information can then be conducted by physicians based on TCM theory and clinical experience to reveal the aetiology (syndrome) of diseases, and thus guiding individualised herbal prescriptions. Though huge leaps have been seen in past decades in elaborating the pathogenesis of CAG, modern medicines remain unsatisfied in treating CNAG and CAG to some extent.9 Owing to the heavy burden of recurrence, chronicity and malignant transformation of CG, a considerable proportion of sufferers in China have focused on TCM therapies.

Trials have indicated the benefits of TCM syndrome-targeted interventions on CNAG and CAG.10-14 The Society of Gastroenterology, China Association of Chinese Medicine has propagated the latest ‘TCM Consensus on Chronic Superficial Gastritis Diagnosis and Treatment’ and ‘TCM Consensus on Chronic Atrophic Gastritis Diagnosis and Treatment’ respectively in 2009.15 16 The Society of Gastroenterology, China Society of Integrated Traditional Chinese and Western Medicine also issued ‘Integrative Medicine Consensus on Chronic Gastritis Diagnosis and Treatment in China’ in 2012.17 The aforementioned consensuses emphasised the importance of applying TCM syndrome-targeted therapies in controlling gastric mucosal inflammation and atrophy, and inhibiting metaplasia and dysplasia in accordance with China’s local conditions. ‘Consensus on chronic gastritis in China’ promulgated by the Society of Gastroenterology, Chinese Medical Association and ‘Diagnosis and Treatment Guideline for Gastric Carcinoma’ issued by the National Health and Family Planning Commission of the People’s Republic of China also recommended the application of TCM syndrome differentiation-based therapies in addition to conventional treatment respectively.1 18

Both consensus and literature reviews have also indicated the potential links between syndrome patterns and CG malignant transformation.19-22 We believe that the diversity and dynamics of TCM syndromes may play a role all the way through this course. However, no large population-based study has been reported to support in-depth analysis of prevalence and severity of varied TCM syndromes in different stages of CG malignant transformation. Taking note of the limitations of previous studies, we highlight validating detailed characteristics of TCM syndromes in CNAG, CAG and GC population respectively in the present study. The new findings may shed light on the TCM pathogenesis theory of CG malignant transformation, and thus optimise syndrome-targeted treatment strategies by TCM herbal prescriptions in distinct stages of this process.

STUDY AIMS

The primary goal for the present study is to identify TCM syndrome features in different stages of CG malignant transformation respectively, including CNAG, CAG and GC. In addition, this study can also be regarded as a pilot study speculating potential dynamic features of TCM syndrome in the whole course of CG. These features may provide important clues for the design of future observational long-term follow-up studies aiming to further explore novel risk prediction and syndrome evolution models combining both TCM and modern medicine indicators for CG or CAG malignant transformation.

METHODS AND ANALYSIS

Design

This is a cross-sectional study undertaken from 2014 to 2018, comprising 2000 eligible participants in four hospitals in China. The general flow diagram of this study is outlined in figure 1.

Recruitment

This study is led by Beijing University of Chinese Medicine (BUCM). The four medical centres participating in the present study include Dongzhimen Hospital Affiliated to BUCM, Dongfang Hospital Affiliated to BUCM, the Third Affiliated Hospital of BUCM, and Wangjing Hospital Affiliated to China Academy of Chinese Medical Sciences. To make a definitive diagnosis for each participant, an upper gastrointestinal endoscopy-based screening process is conducted by study personnel during recruitment. All the eligible patients presenting to the aforementioned hospitals are then identified and recruited continuously until the required sample size for each category (CNAG, CAG and GC) is achieved respectively. This study has opened to recruitment in 2014. Recruitment is ongoing and is expected to conclude by the end of 2018. The sample size required is calculated using the following formula: N=(Z^2 * p(1−p))/d^2, where N is the sample size; P is expected proportion of certain syndrome, if 50%, p=0.5; d is precision, if 5%, d=0.05, which is conventional; Z statistic (Z) is the level of confidence of 95%, conventionally, Z value is 1.96 for 95% CI. According to this formula, the sample size would be 385 cases for each category (CNAG, CAG and GC). However, taking into full consideration of items of case report
forms (CRF), experts’ opinions, study expenses and operability, the anticipated sample size is finally decided to be 2000 in total, including 500 CNAG cases, 1000 CAG cases and 500 GC cases.

**Diagnostic criteria**

Diagnostic criteria of CNAG, CAG, gastric mucosal atrophy (GMA), intestinal metaplasia (IM) and gastric intraepithelial dysplasia (GED) in the present study refer to ‘Consensus on chronic gastritis in China’ promulgated by the Society of Gastroenterology, Chinese Medical Association. Diagnostic criteria of GC refer to ‘Diagnosis and Treatment Guideline for Gastric Carcinoma (trial version)’ issued by the National Health and Family Planning Commission of the People’s Republic of China. Diagnostic criteria of *Helicobacter pylori* (Hp) infection refer to ‘Chinese Consensus on *Helicobacter pylori* Infection Treatment’ issued by the Chinese Society of Gastroenterology, Chinese Medical Association. Criteria for TCM syndrome pattern differentiation diagnosis in this study refer to ‘TCM Consensus on Chronic Superficial Gastritis Diagnosis and Treatment’ and ‘TCM Consensus on Chronic Atrophic Gastritis Diagnosis and Treatment’ issued by the Society of Gastroenterology, China Association of Chinese Medicine.

**Inclusion criteria**

The inclusion criteria are as follows: (1) meeting diagnostic criteria of CNAG, CAG or GC; (2) willing to cooperate with investigators for data and tissue sample collection; and (3) willing to sign informed consent.

**Exclusion criteria**

The exclusion criteria are as follows: (1) history of previous stomach surgeries; and (2) unable to participate in data and sample collection for any reason.

**DATA AND SAMPLE COLLECTION**

Strict training in standard operating procedure (SOP) of clinical data and sample collection for research personnel is conducted by principal investigator before study initiation. Prior to recruitment, participants are required to sign informed consents. In this process, clinical investigators take the responsibility for explaining the objectives, general procedures, data collection methods, risks and benefits, authorisation, data privacy policy and other necessary details to every patient. At the time of the data collection the investigators are blinded for the distinct groups that patients may be included. Research personnel will then assess all the indicators and tissue samples needed at once during recruitment. Each participant is interviewed by at least two physicians to ensure quality control. A list of degree definition of all the TCM indicators is attached to each CRF for investigators and participants to refer to.

**OUTCOME MEASURES AND OVERVIEW OF CRFS**

Primary outcome measures of this study include the prevalence of TCM syndrome patterns in varied stages of CG malignant transformation (CNAG, CAG and GC). Secondary outcome measures include prevalence and severity of all the presenting signs and symptoms collected by using TCM four diagnostic methods (looking, listening/smelling, asking and palpating). The aforementioned clinical features are identified in detail based on TCM syndrome pattern differentiation scales. Accurate modern medicine diagnostic information in this study is achieved based on latest gastroscope reports and pathological reports. Hp infects more than half of the populations worldwide and causes long-term progressive damage to the gastric mucosa in all infected individuals, and is acknowledged as one of the most important risk factors for CG and GC. Stage-specific dietary intake,
and dysplasia, CAG with/without Hp infection, and number of specific populations (eg, CAG with metaplasia included. Subgroup analysis can thus be carried out in a...

According to study design and CRFs, detailed content of data and tissue sample collection are illustrated in table 1.

### PATIENT AND PUBLIC INVOLVEMENT

Before the formal recruitment started, a pretest of CRF was conducted by research personnel in patients at the Dongzhimen Hospital Affiliated to BUCM. In this process, principal investigator and research personnel collated the feedback from the patients and used it to improve the final design of the CRF. Study outcomes will be disseminated on conference reports and academic publications. Main findings will also be disseminated to study participants via email and the BUCM website.

### DATA MANAGEMENT

SOP of data collection, entry, locking and retrieval is set up by data management centre of this study in BUCM. Confidentiality, authenticity and integrity of all the clinical information gathered in the present study are highlighted and maintained at all levels of data management. CRFs are organised by unique identification numbers and thus no patients’ personal identifiers will be available in any records. Before data entry, double-checking of range and logic of every variable value is performed by primary researchers. Data are then double-entered into a dedicated database by data administrators to confirm there is no input error.

### DATA ANALYSIS AND OUTCOME REPORTING

In order to achieve the primary goal of validating TCM syndrome features in varied stages of CG malignant transformation, varied data analysis strategies on multidimensional indicators will be performed parallel from both biostatistics and data mining perspectives.

Descriptive statistics will be carried out by biostatisticians for all the measurement data, including demographic and clinical characteristics in the whole study population and each stratum. Comparative analysis of qualitative TCM and modern medicine indicators in different groups will be performed using X² tests. Comparative analysis of quantitative scores of TCM syndrome pattern differentiation scales will be conducted using rank-sum test. A probability of p<0.05 will be considered statistically significant for these analyses. Correlation analysis of pathological grades and stages of CG malignant transformation and TCM indicators will be performed using logistic regression model. The significance level for introducing and removing variables will be 0.05 and 0.10.

In addition, unsupervised data mining analyses, including exploratory factor analysis, association rule analysis, hierarchical clustering analysis, complex system entropy clustering analysis, and so on, will be performed by data scientists respectively for in-depth exploring of

---

**Table 1 Content of data and tissue sample collection**

<table>
<thead>
<tr>
<th>Content</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demographic information</td>
<td>Name, date of birth, gender, age, nationality, height, weight, level of education, and so on</td>
</tr>
<tr>
<td>Diagnostic information</td>
<td>Latest gastroscope reports, Hp test reports and pathological reports</td>
</tr>
<tr>
<td>TCM syndrome-related information</td>
<td>Features of presenting signs and symptoms collected by using TCM four diagnostic methods of looking, listening/smelling, asking and palpating TCM syndrome pattern differentiation scale scores Results of TCM syndrome identification</td>
</tr>
<tr>
<td>Disease history</td>
<td>Patient history and family history of CG, GC and Hp infection, and so on</td>
</tr>
<tr>
<td>Complications</td>
<td>Upper gastrointestinal haemorrhage, upper gastrointestinal ulcer, pyloric obstruction, anaemia, gastric carcinoma metastasis, and so on</td>
</tr>
<tr>
<td>Dietary intake and lifestyle</td>
<td>Dietary structure, working status, psychological status, history of smoking and alcohol intake, and so on</td>
</tr>
<tr>
<td>Medication</td>
<td>Current drug use on CG and GC treatment (current use of proton pump inhibitors, and so on, at the time of upper endoscopy), and non-steroidal anti-inflammatory drugs (NSAID)</td>
</tr>
<tr>
<td>Laboratory indices</td>
<td>Routine blood test, routine stool test, tumour markers, and so on</td>
</tr>
<tr>
<td>Tissue samples</td>
<td>Tissue samples of diseased gastric mucosa</td>
</tr>
</tbody>
</table>

CG, chronic gastritis; GC, gastric carcinoma; Hp, *Helicobacter pylori*; TCM, traditional Chinese medicine.
TCM syndrome distribution and evolution features in the whole course of CG malignant transformation.

Statistical analyses in this study will be carried out using SAS software (V.9.3, SAS Institute) and R software (V.3.3.1, The R Project for Statistical Computing, Auckland, NZ). Data mining analyses in this study will be processed by SPSS Clementine software (V.12.0, SPSS) and MATLAB software (V.2016a, MathWorks, Natick, MA, USA). Plotting in this study will be performed using Office Excel software (V.2007, Microsoft, Redmond, WA, USA) and Office Visio software (V.2007, Microsoft).

General flow diagram of biostatistical analysis process is outlined in figure 2. Detailed flow diagrams of aforementioned exploratory data mining strategies to be conducted in this study are illustrated respectively in figures 3–5.

**Figure 2** Flow diagram of statistical analysis strategy in this study. TCM, traditional Chinese medicine.

**Figure 3** Flow diagram of exploratory factor analysis in this study. TCM, traditional Chinese medicine.

**STUDY REGISTRATION**
This study has been registered on ClinicalTrials.gov (No NCT03314038).

**ETHICS AND DISSEMINATION**
The protocol is designed in accordance with the Biological Ethics Review Method Involving Humans by the Ministry of Health of the People’s Republic of China, and the Declaration of Helsinki. Ethical Review Board of Dongzhimen Hospital Affiliated to BUCM has approved the protocol (No ECPJ-BDY-2014-02). All the information will be used only for academic aims. All the study outcomes will be disseminated on national conference reports, peer-reviewed TCM and modern medicine journals, or other academic publications. Findings from this study will provide much needed understandings in this topic for both TCM clinical practitioners and researchers, and greatly contribute to optimisation of further prospective study design.

**STRENGTHS AND LIMITATIONS OF THIS STUDY**
Strengths of our study should be addressed. First, the current large cross-sectional study is the first research conducted to reveal TCM syndrome features in varied stages of CG malignant transformation. Although the current study design may limit generalisability of the findings, considering the expenses of prospective study and the relatively low yearly malignant transformation rate of CG individuals, it is important for us to carry out this pilot study speculating potential TCM syndrome dynamic features from CG to GC. The study outcomes will provide targeted clues for the future design of long-term follow-up studies aiming to further explore novel risk prediction and TCM syndrome evolution models for CNAG or CAG malignant transformation combining both TCM and modern medicine indicators. Second, diagnosis of
CNAG, CAG, GC, IM, GED and Hp infection in this study is confirmed based on latest gastroscope reports, Hp test reports and pathological reports at the time of recruitment. The accurate diagnostic information collected will greatly support conducting detailed subgroup analysis stratified by specific stages and pathological conditions (eg, CAG with IM and/or GED, CAG with/without Hp infection, and so on). Third, full-scale information collected in this study provides a good basis for in-depth exploration of clinical features. In addition, the parallel applied biostatistical analysis strategies and unsupervised data mining strategies can reinforce complementary strengths and help us achieve a better understanding of TCM syndrome features in varied stages of these diseases.

Our study also has several limitations. First, this cross-sectional study merely involves data collection at a defined time (recruitment), dynamic characteristics of each individual in the whole course of disease could thus not be tracked. Further large prospective studies with long follow-up time are highly demanded to expand findings of this study in a broadened...
range of settings, and thus contribute to building longitudinal data-based novel risk assessment models of CG malignant transformation combining both TCM and modern medicine indicators. Second, as data are collected merely from TCM hospitals in Beijing, the current participants may not be representative enough to show the clinical features of patients from modern medicine hospitals, or medical centres in other regions of China. Selection bias may thus exist. Third, some of the signs and symptoms-related TCM indicators may be complicated enough for patients to fully understand and correctly report without prior TCM knowledge. A list of degree definition of all the TCM indicators has been attached to each CRF for investigators and participants to refer to in recruitment, and each patient is interviewed by at least two TCM physicians to ensure quality control in this study. However, information bias may still inevitably exist to some extent, especially due to the complexity of individual self-reported TCM indicators.

**References**