General Introduction
Chinese medicine has been used for several thousand years to treat and prevent illnesses. Its superior clinical therapeutic efficacy has won universal praise. Today, as natural methods of healing are promoted, Chinese medicine is receiving worldwide attention. Ever since Chinese medicine was used in the treatment and prevention of illnesses, problems with the authenticity and quality assessment of Chinese medicine have also arisen. Over time, the methods, techniques and theories of assessing the quality of Chinese medicine have also gone through the phases of formation, development and continuous improvement. In the beginning, identification of Chinese medicine was mainly based on the personal experience of individuals. Later, this gradually developed into four major identification methods, namely, origin identification, traits identification, microscopic identification, and physical and chemical identification. Subsequently, many modern identification methods based on instrumental analysis appeared. As modern analytical techniques progressed by leaps and bounds, and research into the Chinese medicine system became increasingly established, the fingerprinting technique for quality control in Chinese medicine was developed. This has accelerated the modernization process of quality control in Chinese medicine.

Here, a fingerprint refers to the common peaks of chromatograms which can be used to characterize a Chinese medicine and are obtained through specific analytical methods. A fingerprint reflects the overall nature of a Chinese medicine, which elevates the quality control target from analysis of a single component to analysis of all the material in a
Chinese medicine. This ensures consistency in the clinical therapeutic efficacy of Chinese medicine, and makes research in this field more aligned with the holistic approach of Chinese medical science. Hence, this has been considered a milestone in Chinese medicine quality control. In recent years, fingerprinting technology has been increasingly applied in many areas such as quality assessment, resource development and the search for active ingredients. It has become one of the important means for variant identification and quality assessment of Chinese medicine. The establishment of a Chinese medicine fingerprint database provides a reference and guideline for routine testing and analytical work, and serves as a solid foundation for the quality standards of Chinese medicine.

**Origin and Development of Fingerprinting**

**Concept of Fingerprinting**

The fundamental difference between Chinese medicine and chemically synthesized drugs is that the effectiveness of a Chinese medicine depends on how various chemical components work together, while that of a chemically synthesized drug usually depends on only one chemical component. So the practice of evaluating traditional Chinese medicine based on the quantitative and qualitative analysis of a single chemical component is gradually being questioned in terms of its effectiveness and accuracy. This is because it is difficult to effectively evaluate the quality and authenticity of a Chinese medicine based on any single active ingredient. This is especially so if the analyzed ingredient is also a common ingredient in many Chinese medicines, thus further reducing the degree of accuracy and specificity of the identification process. In order to increase the accuracy and objectiveness of the identification process, modern analytical techniques are used to analyze the overall effective components of different herbs. One of the techniques, fingerprinting, is based on this concept and has been attracting more interest in recent years.
Fingerprinting is a quality control model that builds upon spectroscopic and chromatographic technology. It is different from the traditional quality control model in the sense that fingerprinting looks at the “complete information” or comprehensiveness of the chromatograph, and displays integrated quality information. Since the secondary metabolites, which are chemical components of medicinal herbs, are inherently unstable, the fingerprints of these chemicals possess a fuzziness that cannot be precisely measured, just like the fuzzy phenomenon in our daily lives. Comprehensiveness and fuzziness are the two basic traits of a fingerprint. The similarity of fingerprints is established through these basic traits. Fingerprint analysis focuses on accurate identification (of similar peaks), and not on precise calculation. The comparison of fingerprints emphasizes similarity and the fingerprints compared do not need to be exactly the same. When it is impossible to find out all the complex components of a Chinese medicine, fingerprints can be used to check the stability of the intrinsic quality of the medicine.

In summary, fingerprinting can be defined as a type of comprehensive and quantifiable method for authentication; it is one of the quality control models that are suitable for evaluating the authenticity and stability of Chinese medicine. Comprehensiveness and fuzziness are its basic traits. The fingerprints should be specific and reproducible and the techniques of the fingerprinting should be practical. This definition explains the essence of Chinese medicine fingerprinting. It shows that fingerprinting is an integrated and comprehensive method of identification. Not only does it differentiate authentic medicines from fakes, it can also evaluate the stability of the quality of a Chinese medicine. Based on its integrative and comprehensive nature, it can be used for accurate and quantifiable authentication and quality evaluation of Chinese medicine. Yet from the perspective of details, the results are fuzzy. Especially since Chinese medicinal herbs are influenced by factors such as the ecological environment and the time of collection, it is very difficult to provide a ratio of common peak to non-common peak that is applicable to all the test samples.
Evolution and Current Status of Fingerprinting

Evolution of fingerprinting

The Chinese medicine fingerprinting technique originates from the development of modern analytical techniques. In the 1970s, trials were carried out to use thin-layer chromatography (TLC) to analyze Chinese medicine. Due to the limitations of the conditions, a lack of maturity in the technique and the timing, it did not receive public acceptance and disappeared from the scene after a period of time. Thereafter, studies in this area were mainly restricted to academic research. In the traditional quality control model of Chinese medicine, there was no requirement for fingerprinting. Hence, it made no impact on the Chinese medicine manufacturing industry. Ever since the State Food and Drug Administration of China enforced the use of fingerprinting as a quality control method for Chinese medicine injection, it has drawn keen nation-wide attention from departments involved in Chinese medicine production, teaching, research and quality control. This generated a strong driving force for research in Chinese medicine fingerprinting.

Fingerprint analysis is an extension of the traditional authentication techniques. In the 1985 edition of *Pharmacopoeia of the People’s Republic of China* (Part 1), besides authentication through physical and chemical methods, there was also a greater use of TLC for authentication. During that time, TLC authentication was all conducted in relation to chemical references, and the requirement was met when the sample showed the same spot as the reference material in the chromatograph. However, due to the limited number of chemical references available, authentication could not be carried out when such references were absent. In addition, to rely on one chemical reference only is not specific enough, especially for a chemical component that is common in many medicinal herbs. Hence, when the 1990 edition of *Pharmacopoeia of the People’s Republic of China* was released, herb references were added for TLC authentication. This resolved to a certain degree the pressing problems of limited availability of chemical references and the non-specificity of using only one chemical reference. In addition, a herb reference in a chromatograph provides much
more information than a single chemical reference. The change or addition of herb references, quantification of samples, comparison of whole chromatographs, and authentication of test samples based on whole chromatographs all possess the initial structure of fingerprinting. However, the process leaves room for improvement due to a lack of standardization of the test samples, test equipment and test procedures. And because of that, the concept of fingerprinting was also not established during that time.

Since the initial trials in the 1970s, and with the development of modern analytical techniques, subsequent academic research, the foundation of the widespread application of herb references for TLC authentication established in the 1990, 1995 and 2000 editions of *Pharmacopoeia of the People’s Republic of China*, we have taken a step further in the standardization of medicinal materials, operating conditions and test methods. The concept of Chinese medicine fingerprinting has come into focus and should gradually develop as an extension of current Chinese medicine authentication. Also, through fingerprint research, the overall standard of Chinese medicine cultivation, collection, processing, production, analysis and quality control is elevated. The diversity and continuous improvement of the level of current instrumental analysis techniques have also provided a good working platform for fingerprinting research and applications in Chinese medicine.

**Current research in fingerprinting**

Fingerprinting has become an important domain and hot spot for current fundamental research in Chinese medicine in China. There are a good number of academic research reports. The content includes theoretical discussions of Chinese medicine fingerprinting, standardized testing for the analysis of fingerprints, research in the methodology of establishing fingerprints, informatization and knowledge discovery research, as well as the applications of fingerprinting in improving quality control standards, guiding standardized production, developing new resources and new formulations of Chinese medicine, etc. The research has obtained rather satisfactory results.
practical applications, the national Chinese medicine injection fingerprint research plan has been launched officially. More than a hundred specialists and scholars from 48 research institutions and institutions of higher learning will complete a study of 74 types of preparation involving 146 Chinese medicine injection production enterprises. The State Food and Drug Administration promulgated the Guidelines for the Test Procedures of Chromatographic Fingerprints of Chinese Medicine Injection and two types of computer-aided Chinese medicine fingerprint similarity computation software. The guidelines detail the requirements for test sample collection and preparation, selection of reference, fingerprinting test conditions, laboratory and instrumentation facilities, test method, and the construction, identification, verification and checking of chromatographic fingerprints for raw materials, medicinal materials, intermediates and final products. The two types of computation software are able to compare and compute all the Chinese medicine chromatographic peaks, to determine the overall degree of similarity between the fingerprints. They automatically process the information to accurately evaluate the stability of the quality of the Chinese medicine. This established a practical technical platform and scientific foundation for the comprehensive research of Chinese medicine fingerprinting. Currently, the study of Chinese medicine fingerprinting is still primarily at the initial stage, which focuses on determining the methods of obtaining fingerprints, determining the evaluation criteria, and evaluating and controlling the quality of Chinese medicinal material.

Outside of China, fingerprinting is listed in the US Food and Drug Administration’s (FDA) Guidance for Industry Botanical Drug Products (drafted in year 2000), the World Health Organization’s (WHO) Guidelines for the Assessment of Herbal Medicine (1996), the British Pharmacopoeia (1986), the Indian Pharmacopoeia (1998), and the US Pharmacopeia (1999–2001). The objectives of these guidelines are to resolve the problem of unknown active constituents in many herbal medicines and to address the problem of authenticating and controlling the quality of commercial products. They require the herbal medicine product manufacturers to provide the fingerprints of intermediates to ensure the authenticity of the species, and the fin-
fingerprints of final products to verify the batch-to-batch stability in quality. The US FDA allows the provision of chromatographic fingerprint authentication information for applications relating to herbal health products. In the WHO’s Guidelines for the Assessment of Herbal Medicine, where the active constituents cannot be distinguished in the herbal medicine products and final products, provision of the fingerprints of characteristic constituents or mixture of constituents will ensure consistency in the product quality. The number of overseas research articles on fingerprinting is also growing. These articles recognize too that the efficacy of a particular herbal medicine cannot be attributed to a single constituent. In some herbs, the active constituent is not known, and is difficult to identify. Hence, fingerprint identification is brought up as a possible solution. Basically, the research and application of fingerprinting is currently restricted to monitoring and ensuring the stability of the inter-batch quality of herbal medicine with complex constituents and unknown active constituents.

Classification and Development of Fingerprinting

Classification of fingerprints

The fingerprint that we refer to presently is an extension of and development from the use of complete chromatograms of reference herbs to identify test samples. A fingerprint, in the narrow sense, refers to the fingerprint of the chemicals in a Chinese medicine, which represents the characteristics of the chemical metabolites in the medicinal herbs. A fingerprint in the general sense can be classified into different types according to the object of application or the method of analysis.

Classification according to object of application

Fingerprints can be used at every stage of the preparation research and manufacturing process. They can be classified into Chinese medicinal material (raw material) fingerprints, processed medicinal material (including sliced medicinal material, blended pellets) fingerprints,
intermediate (intermediary product in the production process) fingerprints and preparation fingerprints, depending on the object of application.

Chinese medicinal material fingerprints can be further classified into chemical fingerprints and DNA fingerprints, according to the method of analysis. Chinese medicinal material DNA fingerprints are mainly for the determination of the DNA profile of each medicinal material. Due to the unique and hereditary nature of the genes of every species, DNA fingerprinting of Chinese medicinal materials can be used in species identification, botanical classification research and quality research. It is extremely useful in research on the setup of Good Agricultural Practices (GAP) bases, cultivation standards (standard operating procedures or SOP), selection of premium cultivation resources and the authenticity of Chinese medicinal herbs. Chinese medicinal material chemical fingerprints refer to fingerprints setup through the analysis of each chemical constituent (secondary metabolites) in the medicinal material. Although chemical constituents are secondary metabolites, which are affected by the biological environment and life span and so produce obvious differences in individual plants, the metabolism of plants is hereditary in nature and plants of the same species will possess similarities in their chemical composition. Hence, the chromatograms of the chemical constituents can be used to setup the fingerprints. The use of chemical fingerprints of Chinese medicinal materials in quality control is more direct and of greater significance.

Processed medicinal material fingerprints should be treated separately. The requirements of the fingerprints for sliced medicinal materials are basically the same as for the medicinal material itself. For blended pellets (including standard extracts), the fingerprint requirements should be greater. The standard extraction procedure should be carried out. Operating parameters should be controlled strictly. The method of mixed-batch blending should be used to process medicinal materials, so that it is completely feasible to establish a stable fingerprint for quality control.

In establishing stable chemical fingerprints for Chinese medicine preparations and intermediates, if one can change the feeding materi-
als from raw materials to extracts (maybe using the mixed-batch blending method), enforce strict control over the process flow, and monitor the entire production process through fingerprinting, the problem of quality stability of the final product could be resolved.

Classification according to method of analysis

Chinese medicine fingerprints can be classified into chemical fingerprints and biological fingerprints according to the method of analysis. Chinese medicine chemical fingerprints refer to fingerprints obtained through chromatographic, spectroscopic and other methods of analysis to establish the characteristics of the chemical constituents. Generally, these include thin-layer chromatography (TLC) fingerprints, gas chromatography (GC) fingerprints, gas chromatography–mass spectrometry (GC-MS) fingerprints, high-performance liquid chromatography (HPLC) fingerprints, high-performance liquid chromatography–mass spectrometry (HPLC-MS) fingerprints, high-performance capillary electrophoresis (HPCE) fingerprints, high-speed counter-current chromatography (HSCCC) fingerprints, supercritical fluid chromatography (SFC) fingerprints, ultraviolet spectroscopy (UV) fingerprints, infrared spectroscopy (IR) fingerprints, fluorescence photometry (FP) fingerprints, nuclear magnetic resonance (NMR) fingerprints, mass spectrometry (MS) fingerprints, X-ray diffraction (XRD) fingerprints, and hyphenated chromatography fingerprints (multi-dimensional multi-data characteristic chromatograms). Besides these, there are electrochemistry fingerprints, electrophoresis (EP) fingerprints, differential thermal analysis (DTA) fingerprints, circular dichroism (CD) fingerprints, trace element fingerprints, X-ray fluorescence spectroscopy (XRF) fingerprints and oscillating fingerprints.

Chinese medicine biological fingerprints include DNA fingerprints, genomic fingerprints and proteomic fingerprints. Chinese medicine genomic and proteomic fingerprints refer to the use of Chinese medicine preparations to act on certain cells or animals, resulting in complex changes in the condition of the genome and proteins. These two types of fingerprints are also called bioactive fingerprints. Besides,
there are scanning electron microscopy (SEM) fingerprints and computer image analysis (CIA) fingerprints.

**Development of fingerprinting**

Fingerprinting, as a breakthrough in the modernization of Chinese medicine, should not be restricted to the elevation of the quality standards of Chinese medicine. Instead, one should get to know fingerprinting from the vantage point of how the modern scientific system should be used with traditional Chinese medicine, and finally to work towards the globalization of Chinese medicine. Modern cross-disciplinary language could be used to describe the essence of Chinese medicine. After decades of research and development, copious research has been conducted on Chinese medicine (Chinese medicinal material, sliced medicinal material and Chinese proprietary medicine), accumulating a large amount of research data on the chemical compositions of single medicinal herbs, research approaches and partial data on the study of chemical compositions of Chinese medicine formulae. The rapid development of modern analytical science has seen the emergence of numerous advanced methodologies, stable and reliable equipment and technology for setting up fingerprints. The rapid development of hyphenated technology has, in particular, provided the possibility of online analysis of unknown components. The development of information science allows the huge amount of data obtained to be informatized. Complex systems science has provided us with an overall approach to and a corresponding method for resolving such a complex system as Chinese medicine. All these helped to establish a solid foundation for the large-scale development of Chinese medicine fingerprint research, for the establishment of standardized and functional fingerprints, and for the integrated evaluation of Chinese medicine.

**Developmental phases of fingerprinting**

The research and construction of Chinese medicine fingerprints can be classified into initial and advanced phases.
The initial phase should determine the fingerprint construction method and assess the degree of similarity. Chinese medicine injection is fast-acting, has good therapeutic effects, and its chemical constituents are comparatively fewer than other forms of Chinese medicine. Yet the requirements for product quality are more stringent compared to those of other dosage forms. It is highly feasible to choose Chinese medicine injection to proceed with the work of setting up fingerprints. In the case of Chinese medicine injection, for the initial phase, the fingerprints are analyzed through the injection forms (final products), intermediate products and raw medicinal materials. The corresponding indicative control parameters are setup, to achieve the aim of obtaining operable, controllable, stable and quantifiable fingerprints. Fingerprints can be used to represent the final product quality, to establish quality control for the production process and the medicinal materials, to trace the source of problems in the production process and to demonstrate the quality requirement for the GAP of the medicinal herb. After a large amount of fingerprint data has been obtained, the method and indicators for determining the degree of similarity among fingerprints can be discussed. One can use the “common peak” method to determine the degree of similarity for the purpose of control, or use pattern recognition to determine the degree of similarity among the overall fingerprint chromatograms, and determine the various control parameters. Research in the initial phase focuses on setting up the analysis method, investigating the method stability and suitability, basic research on the active chemical components, multi-indicator control of the fingerprint, etc, which elucidate the chemical basis of the Chinese medicine and the quality control parameters, and makes quality control of Chinese medicine quantifiable.

After a relatively complete Chinese medicine fingerprint has been setup, one enters the advanced phase of the research, which is the investigation into the relationship between fingerprint characteristics and drug efficacy, and the bioequivalence of fingerprints. This is the basic meaning of multi-dimensional and multi-data characteristic fingerprints and fingerprint pharmacodynamics. Western medicinal compounds have a quantitative structure-activity relationship, which means a certain molecular structure has a corresponding biological...
activity. Chinese medicine takes effect through a group of chemical compounds that have drug efficacy. Different ways of herb blending (including changes in the type and quantity of herbs) will result in changes in the type, number and quantity of chemical compounds. Such changes in chemical groups can be expressed through fingerprints. The correlation between the changes in chemical composition (type, number and quantity) and drug efficacy (pharmacodynamic studies or clinical effects) can be established through a fingerprint. At this point, the fingerprint obtained is not only the integral expression of the chemical composition; it also contains biological information on related drug efficacy. Hence, one is able to achieve corresponding drug efficacy and clinical effects though control of the fingerprint. For different Chinese medicinal herbs, the differences in fingerprints are even greater; the corresponding drug efficacies are also not the same. By making changes to the drug blending and conducting research on the quantitative correlation between chemical composition and drug efficacy, one is able to use fingerprints to express different biological information. If new drug development is able to make use of fingerprints, right from the start, to manifest the changes in the chemical groups, together with pharmacological and pharmacodynamic testings as well as correlation analysis, the development of new drugs will be sped up and a higher standard can be obtained.

There is no clear demarcation between the initial and the advanced phases. In the initial phase, the main objective is to carry out relatively comprehensive quality control. Hence, the research should be in-depth, and during execution, it should be simplified so as to facilitate its widespread use. The data accumulated during the initial phase provides the foundation for research in the advanced phase. Fingerprints that are able to express integrated information about the correlation between chemistry and drug efficacy will help to speed up the modernization of Chinese medicine, and will provide reliable quality assurance for the globalization of Chinese medicine.

Informatization and knowledge discovery of fingerprints

In the development of Chinese medicine fingerprinting, the advancement of information science can help the informatization and
knowledge discovery of fingerprints. Fingerprint informatization includes data acquisition and digitization. The main process includes selecting a suitable analysis method to obtain the fingerprints, setting up the operating conditions for the entire analysis procedure, and conducting tests to obtain the fingerprints of various samples, which is known as data acquisition. The obtained fingerprints are analyzed to determine their numerical characteristics. For example, in infrared fingerprints, the peak numbers and peak intensities can be the numerical characteristics used to define the peaks. For HPLC fingerprints, the peak parameters such as the retention time (or relative retention time) and peak height (and/or peak area) can be the numerical characteristics. Direct observation of the actual fingerprints allows easy comparisons between the fingerprints under simple conditions, whereas extraction of the numerical characteristics to obtain a digitized fingerprint allows for ease of mass storage, comparison and utilization. Hence, a database should be setup for large volumes of fingerprints. For Chinese medicine, a fingerprint database for different variants, places of origin, times of collection and processing methods should be setup for ease of comparison, so as to determine the degree of variation. As for Chinese medicine formulation (Chinese proprietary medicine), besides a fingerprint database for the Chinese medicine formulation, a fingerprint database of the entire process including raw materials, intermediates or effective parts, and final products should be setup as well. This is to facilitate quality control in the entire process and to trace the effect of changes in operating conditions and raw materials on product quality. For new drug research and development, the fingerprints of research samples at each phase can be collected into a database, which will enable monitoring and characterization of the chemistry and drug efficacy of each effective part. The capacity to add methods and functions to evaluate and compare fingerprints to those in the existing database will render it an intelligent database. A sample fingerprint could then be keyed in and compared with existing fingerprints, and hence interpreted to determine its quality.

Fingerprint knowledge discovery refers to the study of how to utilize the information after the fingerprints have been informatized. Fingerprint knowledge discovery includes interpretation, comparison
and evaluation of information, study of the correlation between chemistry and drug efficacy, and information utilization, which is the extraction of regular patterns and knowledge from a large volume of fingerprint data. China has launched some research work in this area. Chinese medicinal materials have been extracted and separated to obtain the constituent chemicals. Under the guidance of traditional Chinese medicine (TCM) theory, stepwise regression analysis and canonical correlation analysis are conducted on pharmacological and chemical data obtained from related pharmacological experiments, to confirm that a particular chemical constituent in the Chinese medicinal material is the main material basis for its effectiveness. The particular active constituent is chosen as the subject for study. The sample medicinal materials are extracted and separated to prepare sample solutions, and the fingerprints (such as HPLC fingerprints) of the active constituent are analyzed. The chromatographic peaks are obtained and used as the variables for chemical characteristics. Reference peaks are also used in fingerprint analysis. After the fingerprints are informatized, chemical pattern recognition is carried out to determine the quality of the medicinal materials. In the case of Chinese medicine preparation, based on the known therapeutic effects of the prescription, the corresponding pharmacodynamic test method is determined. Based on the current chemical and pharmacological information on the medicinal materials (or effective parts and constituents) that make up Chinese proprietary medicine, and production technology, the possible class of chemical constituents is determined. Modern separation techniques (such as supercritical fluid extraction, macroporous resin adsorption, counter-current chromatography and all types of preparative chromatography) are used to separate the sample into each chemical type (chemical group) for pharmacodynamic testing, in order to finally determine the effective parts (the chemical constituents that are verified by pharmacodynamic testing) in the Chinese medicine. Its fingerprint (such as a HPLC fingerprint) is established and the chemical information (including the number, content and ratio of chemical constituents) and drug efficacy information obtained are analyzed for overall correlation, to obtain the cause and effect relationship (namely
the regular pattern and knowledge).

An intelligent Chinese medicine fingerprint database is a database that collates a large number of fingerprints from medicinal materials, intermediates and final products. The correlation of the peaks can be determined using mathematical comparison methods. This is such that total quality management can be realized for the entire production process. As for basic research and new drug development, the characterization of chemical constituents for the entire process can be actualized. In addition to multi-dimensional (HPLC-PDAD-MS/MS) fingerprints, a large amount of information needs to be processed for the entire research. Hence, setting up an intelligent fingerprint database will help to resolve this problem.

Developmental prospect of fingerprinting

The development trend of Chinese medicine fingerprinting is built upon its fingerprint pharmacodynamics, actualizing the informatization and knowledge discovery, making Chinese medicine safe, effective and controllable. In actual applications, there are several aspects in Chinese medicine fingerprint quality control technology that can be further developed:

– a rational, easy-to-implement fingerprint similarity analysis method and the corresponding software, that is packaged with the analytical instrument, to be used by Chinese medicine manufacturing companies and drug analysis departments for quality control and evaluation;

– Chinese medicine analysis, identification, classification and evaluation software and network platform to service all the relevant departments, companies and organizations;

– online fingerprint analysis technology, analysis software and production automation based on fingerprint information, so as to actualize the control of Chinese medicine production process using fingerprints;

– a fingerprint database and network platform for Chinese medicinal materials (raw materials, especially for authentic medicinal materials),
processed medicinal materials (including sliced medicinal materials, blended pellets), intermediates (intermediate products from the production process) and preparations, to ensure total quality management of Chinese medicine;

– analysis and evaluation software for the correlation of fingerprint characteristics and drug efficacy, in order to establish fingerprints with integrated information on the correlation between chemistry and drug efficacy, speed up the new drug development process and raise the standard of research.

In summary, Chinese medicine fingerprinting possesses promising developmental prospects and potential for important applications. It will definitely play an even more significant role in the modernization of Chinese medicine.

A schematic diagram of the origin and development of Chinese medicine fingerprinting is as shown in Fig. 1.

**Significance and Applications of Fingerprinting**

To date, the quality control technology for Chinese medicine fingerprinting has become vital in driving the overall improvement of this industry. Its application and research have been significant in strengthening the Good Laboratory Practice (GLP) of Chinese medicine, ensuring its efficacy, realizing the Good Manufacturing Practice (GMP) of Chinese medicine, raising the overall standard of the Chinese medicine industry, implementing the GAP of Chinese medicinal materials, leading the agriculture of Chinese medicine towards modernization, and driving Chinese medicine towards globalization.

**Adoption of Fingerprinting Technology to Ensure the Efficacy and Modernization of Chinese Medicine**

Chemical substances in Chinese medicine are very complex. Ginseng is the species with the most comprehensive and the longest history of research on its chemical composition. Its documentation can be traced back to 1854 in a report on ginseng methaqualone by Sarriquex S. in *Annalen der Chemie und Pharmacie*. In the past 30 years, new gin-
Analyzing direct characteristics of Chinese medicine (appearance, colour, taste)

Lab test (visual chemical analysis)

TLC visual analysis for medicinal materials or their controls

TLCs → HPTLC

HPLC-MS → HPLC-NMR

HPLC → HPLC-DAD

HPLC-ELSD

GS-MS → GC

GC → PGC

GC-FTIR

HPCE

HSCCC

CD

SFC

FS → UV → DS

NMR → FT-NMR

MS

NIR

NIRDRS

NIRDRS-CT

IR

FTIR

2D FTIR

DR-FTIR

ATR-FTIR

FT-Raman

XRD fingerprint

DTA fingerprint

DNA fingerprint

Trace element fingerprint

Electrochemistry fingerprint

Electrophoresis fingerprint

Oscillating fingerprint

Informationization and knowledge discovery

Chinese medicine fingerprints

Multi-dimensional multi-data characteristic chromatogram

Intelligent database of Chinese medicine fingerprints

Fig. 1. Schematic diagram of the origin and development of Chinese medicine fingerprinting.
seng components have constantly been discovered. Currently, the known components include more than 30 types of saponins, 29 types of volatile oils, 15 types of amino acids, 29 types of minerals, 16 types of carbohydrates, 11 types of organic acids, esters, alkaloids, vitamins, sterols and many enzymes. A single medicinal material is so complex in nature, as it is, let alone the chemical components in a medicine formula. Based on this, it would be highly impractical if quality control were to be established on the basis of the study of every single chemical component. The requirement now is to study the entire species under the conditions where the chemical composition is not completely known. The fingerprints of species obtained through modern instrumental analysis, such as chromatography, spectroscopy and mass spectrometry analysis, present this possibility. Similarities in the fingerprints of several species obtained using the same instrument, the same operating procedures, and under the same test conditions reflect similarities in their attributes. Even though one may not understand the composition represented by each characteristic peak in the fingerprint, and may not have a complete knowledge of the chemical composition of the species, this does not influence the assessment of the uniformity of the species. It can provide qualitative analysis, as well as semi-quantitative analysis.

In fact, from the point of view of the development of analytical chemistry, analysis of pure substances is not particularly challenging. Yet many disciplines, such as environmental chemistry, have highlighted the requirement to carry out rapid qualitative and quantitative analysis of complex organic compounds. The trend towards the use of instrumentation for analysis and the more complex chemistry system are two important characteristics of modern analytical chemistry. At the same time, chemometrics has been developing rapidly. From the perspective of modern chemometrics, in the traditional method of data processing in instrumental analysis, such as for spectroscopy, only the peak value is taken; for chromatography, only the peak area is calculated. As a consequence, large amounts of useful information are wasted. Chemometrics, which uses statistics, applied mathematics and computers, is able to extract the maximum amount of qualitative and quantitative chemical information from samples with different charac-
teristics. Research in this area also suggested the classification of the analysis system into white, grey and black multi-component systems according to the extent of prior knowledge of the qualitative composition. Modern analytical chemistry, especially the development of chemometrics, has set the technological foundation for the use of Chinese proprietary medicine fingerprinting for quantitative and qualitative analyses.

We have to control the intrinsic quality and at the same time ensure the efficacy of Chinese medicine. We have come to a historical moment in the development of analytical chemistry, where we are able to resolve the quantitative and qualitative analyses of complex multiple components. Hence, there is a need to make full use of the conditions to resolve the problems encountered, break the deadlock and establish a new era where qualitative and quantitative analyses of Chinese medicine are gradually made possible.

**Fingerprinting as the Critical Technology for the Modernization of the Chinese Medicine Industry**

The development of the Chinese medicine industry follows a modernization process. From improvements of plant areas, plant buildings, technologies and equipment, to innovations in packaging and dosage formulation, it is now time to modernize the quality standards of Chinese proprietary medicine. The modernization of the production of Chinese proprietary medicine will eventually be reflected in the modernization of the controllable quality of the products. The controllability refers to the fact that the quality standard of a product should ensure the stability of the drug efficacy. In addition, the quality of the products should be controllable during the entire manufacturing process. These two points are the most significant problems in the modern production of Chinese proprietary medicine. At present, the existing quality standard is not good enough to control the stability and efficacy of Chinese medicine products, and it is also difficult to ensure product consistency during the manufacturing process. All these problems come from our limited understanding of the different components in a product.
Quality control through fingerprinting is expected to significantly increase the level of controllability, and promote the modernization of a product and its production process. When fingerprinting is applied, the quality of the raw materials will need to be tested by the company and the storage management methods will need to be improved. To ensure consistency in the quality of the final products, the dosages of raw materials should be adjusted so as to make up for the differences in quality in different batches of materials. To ensure consistency of products, production processes should be studied to establish strict control parameters. To ensure consistency of operations and end results, management of the manufacturing processes and personnel must be improved. To meet final product specifications, the quality of intermediates should be monitored; technical strategies should be studied to ensure that material feeding in every step meets specifications. To ensure the consistency of equipment conditions and equipment usage, the equipment models should be carefully selected and equipment maintenance should be improved.

Since there are quality control requirements for fingerprints, we now have the criteria for evaluating the impact of technology and equipment changes on product quality during the manufacturing process. With the understanding that quality comes first, one can boldly experiment with various new technologies and equipment.

Fingerprinting technology is closely related to manufacturing procedures. There are many variable factors in this technology, which will cause inconsistencies in the fingerprints, yet the many adjustable factors broaden the field of research. It is similar to the difference between fluidized bed granulation and mechanical granulation, or that between spray drying and box drying. Fluidized bed granulation and spray drying present a lot of variables and are not easily mastered. But although expertise is required for these methods, if they are mastered by the operator, manipulations become very flexible. The industry of Chinese proprietary medicine also needs to overcome some technical and management obstacles when using fingerprinting technology. At the same time, a lot of technical secrets are hidden in the construction of the fingerprint of a new product. In the past, Chinese proprietary medicine formulae were trade secrets. Now, even if a formula is publicized,
without the technical know-how to construct a qualified fingerprint, the formula is of no use. Fingerprints have become technical barriers to protect a company’s new products. Unlike patent protection, this kind of barrier is not restricted by any time limit.

The use of fingerprinting technology for quality control will not only promote the application of new technologies and equipment, but will also open a new channel for research into new Chinese proprietary medicine. The effectiveness of various technologies can be compared by studying the different chromatograms obtained using the same medicine formula but with different production technologies, so that the most optimized technology can be selected.

The difficulty of fingerprinting technology will certainly drive the centralization and standardization of Chinese proprietary medicine production. It plays a guiding role in the modernization of Chinese proprietary medicine.

_Fingerprint Quality Control as a Technical Link to Drive the Modernization of Chinese Medicine_

Chinese agriculture has entered a new phase. While maintaining a stable increase in yield, it is paying more attention to increasing product quality, modifying and optimizing the plant species, and producing high-quality agricultural products. The government has announced guidelines to greatly increase the overall quality and efficiency of agriculture through science and technology. The government has also announced guidelines relating to the mode of production, saying that “industrialization should become an important approach to advance agricultural modernization, to encourage and support agricultural products processing and marketing companies to help farmers enter the market, and finally form a profit- and risk-sharing organization”.

It is also time for the agriculture of Chinese medicine to pay attention to quality and authenticity. How much attention will be paid to quality in the agriculture of Chinese medicine depends on how much attention will be paid to the quality of raw materials by the Chinese proprietary medicine companies. The Chinese medicine industry has been making great progress in the recent 30 years. The weakest link is

---

the purchase of raw materials. Some Chinese medicine companies do not pay attention to the quality of raw materials. They purchase whatever is cheapest, resulting in a situation where attention is paid to quantity, not quality, in the cultivation of Chinese medicine. Medicinal materials, decoction pieces and proprietary medicine are the three mainstays of the Chinese medicine industry. Among them, Chinese proprietary medicine has the highest market share. Seventy percent of the medicinal materials are sold to Chinese proprietary medicine companies. Being the main player in the Chinese medicine industry, the Chinese proprietary medicine industry now plays a leading role in Chinese medicine agriculture.

For a long time, many Chinese proprietary medicine companies were content with using raw materials that merely complied with quality requirements. This was partly a result of paying too much attention to profits, and partly because the quality standards of Chinese proprietary medicine lack specifications on the quality of raw materials. To apply fingerprinting technology in quality control, the quality of the raw materials must be stable. To ensure the quality of the final products, the companies of Chinese proprietary medicine must pay attention to the quality of medicinal raw materials and play a leading role in Chinese medicine agriculture. This kind of leadership role focuses on three aspects. The first aspect is the quality requirements of raw medicinal materials. Most farmers know that Chinese medicine companies need genuine and high-quality medicinal materials. The second aspect is to lead or guide the farmers towards industrialization. The Chinese proprietary medicine industry should pay more attention to helping farmers establish raw material bases and should consider the raw material bases as their first workshop. This is such that the quality of the final products is built upon the high quality and stability of the raw medicinal materials. The third aspect is to begin research into the cultivation techniques used for raw medicinal materials. This means bringing in agricultural technology, selecting better places of origin, and setting up the SOP for producing medicinal materials. In summary, the Chinese proprietary medicine industry should play a leading role in promoting Chinese medicine agriculture to achieve GAP. For major raw materials, companies should put effort
into establishing bases that meet the GAP requirements.

The selection of places of origin of medicinal materials has great significance in base construction, development of SOP for medicinal material production, and application of GAP. There is much data demonstrating that there are huge differences in the concentrations of components in the same kind of medicinal materials if they have different origins. For many herbs, there could be a difference of more than 10 times in the main component concentrations when we compare herbs with different origins. For example, chlorogenic acid in honeysuckle ranges from 0.45% to 3.42%; icariin in *Herba Epimedii* ranges from 0.14% to 3.00%; astragaloside A in astragalus root ranges from 0.013% to 0.30%. That is why the selection of bases in the country should be based on both authenticity studies and scientific analysis.

Developing an SOP for the production of raw medicinal materials is very important to achieve standardized production and consistent quality. With an SOP, farmers will cultivate medicinal materials according to it, just like production workers manufacture medicinal products according to an SOP. In this case, GAP is similar to GMP. How to organize the farmers’ efforts is a big issue. Some companies already have some successful experiences, such as the *Radix scutellariae* base in Zhangjiakou of Hebei Province setup by the Second Factory of Harbin Chinese Medicine; the *Radix salviae miltiorrhizae* base in Nanluo of Shanxi Province setup by Tianjin Tasly Inc. and the saffron crocus base in the suburbs of Shanghai setup by Shanghai Pharmaceuticals Pte. Ltd. There are various types of organization and it is important to experiment with different types of setup.

Besides setting up bases for cultivated species, research into converting wild species to cultivated species should be conducted actively according to industry needs. Bases should also be setup for successfully converted species. In brief, an industry cannot be established on the grounds of nearly exhausted resources. Building a medicinal material base will bring several advantages to the industry, enterprises and farmers, such as stable supply and demand, stable prices, stable production and stable quality. For the Chinese medicine industry, the industry plays a leading role in its modernization. A modernized
industry will lead to a modernized agriculture. Only then can the sustainable and coordinated modernization of the Chinese medicine industry be ensured. One can say that without the modernization of Chinese medicine agriculture, the modernization of Chinese medicine is impossible. The use of fingerprinting has already pushed some Chinese medicine companies to seriously consider the issues of setting up bases for raw materials, the relationship between the Chinese medicine industry and Chinese medicine agriculture, and the relationship between Chinese medicine agriculture, carrying out national policies and achieving modernization. This shows how important this technology is in driving the Chinese medicine industry and Chinese medicine agriculture to jointly achieve modernization. This is only the beginning. The scaling up of the industry has brought about centralized and standardized (GAP) production of Chinese medicine agriculture, and its effect on Chinese medicine agriculture will become more and more significant.

**Fingerprinting as an Essential Quality Control Technique**

In the modern world, “safe, effective, and controllable” are internationally recognized requirements of medicaments. There is also a growing common understanding in the quality control methods in Chinese proprietary medicine in recent years. The major Japanese companies producing Chinese medicine have been using HPLC fingerprinting for quality control since the 1980s. They produced standard fingerprints using decoctions of medicinal teas prepared according to prescription formulae. The raw materials, formulae and operating procedures used for mass production were strictly controlled to ensure that the final product fingerprints would be consistent with the standard fingerprints. Europe has always valued the therapeutic effects of herbal medicine. The Europeans have also been using fingerprinting to control the quality of herbal medicine. One typical example is the extraction of *Foliium Ginkgo* jointly developed by researchers in Germany and France. During their research, they found that the therapeutic effect of *Foliium Ginkgo* extract comes from the synergistic effects of the substances in the extract. They conducted a long-term
research study and found that isolated lactones or flavones from the extract do not have the same effect as that of the extract as a whole. This finding is consistent with our understanding of the effects of Chinese proprietary medicine. HPLC fingerprinting should be used for the quality control of this kind of “whole materials.” For example, Hypericum perforatum has been used in Germany to treat depression. It was introduced to the United States and was very popular in the US herbal market in 1998. In the US, the content of hypericin was used as a quality indicator. However, some studies in recent years have discovered that this component does not have antidepressant effects. Hence, the fingerprinting technology that has been used for quality control by German companies is regarded as a more reliable method. In recent years, the US FDA has listed HPLC fingerprinting as the quality control method for mixtures of substances in its botanical drugs guidelines. In brief, fingerprinting is internationally recognized as the quality control method for Chinese proprietary medicine and herbal extracts. Therefore, a quality control system using fingerprinting technology should be established according to the unique features of Chinese medicine.

Requirements and Construction of Fingerprints

Requirements of Fingerprints

Chinese medicine fingerprints should satisfy the practical requirements of specificity, reproducibility and practicability.

Regardless of whether it is a Chinese medicinal material, Chinese proprietary medicine or Chinese medicine formula, the fingerprint must be able to display the characteristic, also known as the specificity or uniqueness, of that Chinese medicine (herb or formula). Of course, there is the possibility that one fingerprint is not adequate to represent the full characteristics of a Chinese medicine, and so there may be a need to use several fingerprints to demonstrate the different characteristic profiles of a certain Chinese medicine (herb or formula), in order to form the full picture. But for each fingerprint there should still be the requirement of specificity.
Fingerprinting is mainly used to demonstrate and control the chemical constituents of Chinese medicine as a whole. Thus there is a need for good reproducibility, i.e. for the same sample, under the same operating conditions, the level of reproducibility of the result must be good. Hence, according to the differing requirements, one should select the most suitable analytical method for fingerprint construction. One of the main characteristics of Chinese medicine is its complexity. It involves uncertainty and some fuzziness. These problems affect the reproducibility, and could be eliminated in the data-processing steps (such as in the assessment of degree of similarity) through the use of suitable methods.

The practicability of a fingerprint refers to the choice of different analytical methods according to different applications to achieve the different requirements. For example, for quality control, one should consider the use of conventional facilities or equipment found in the manufacturing enterprise or drug analysis department to setup the corresponding method. Usually, these include photometers, GC and HPLC. In the case of using fingerprinting to study the theory of herb compatibility (配伍) or the development of new medicine, especially those in areas related to chemical composition, pharmacology and drug efficacy, one should consider the use of hyphenated technology such as GC-MS and HPLC-DAD-MS/MS to obtain a large amount of data. This is preferable due to the acquisition of unambiguous results.

**Methods of Constructing Fingerprints**

With the rapid development of contemporary chromatographic analysis and photometric analysis, more and more advanced techniques are applied to the quality analysis of Chinese medicine. These are powerful tools for the construction of the Chinese medicine fingerprints. The commonly used methods of constructing fingerprints are the photometric method, chromatographic method, X-ray diffraction method, thermal analysis method, electrophoresis, hyphenated chromatographic technology, molecular biology technology, electrochemistry method, scanning electron microscope, computer image analysis
and chemical oscillation technology. In actual applications, we can select one or more methods according to our needs.

**Photometry**

Photometry refers to the use of light of a certain wavelength to scan the sample and obtain the characteristic spectra and data. The main methods include ultraviolet (UV) spectrum, infrared (IR) spectrum, fluorescence spectrum (FS), nuclear magnetic resonance (NMR) and mass spectrometry (MS). In addition, there is atomic absorption spectrometry (AAS), inductively coupled plasma (ICP) and X-ray fluorescence (XRF).

**UV spectrum**

UV spectrum is the absorption spectrum that demonstrates the absorption of energy in the UV and visible light regions by electrons in the molecule to produce electronic transitions. In Chinese medicines, some of the constituents possess unsaturated or conjugated double bond structures, which under UV light (200–400 nm) can undergo changes in the atomic or electronic active states of the chemical substances. A portion of the energy will be consumed while the rest is transmitted, passed through a prism or diffraction grating to be separated into a non-continuous spectrum based on the wavelength. Dissimilar substances produce dissimilar UV absorption spectra. Since different Chinese medicines exhibit diversity in the unsaturated components, their UV spectra would also demonstrate characteristic differences (maximal absorption wavelength, minimal absorption wavelength, peak shoulder, absorption coefficient, absorbance ratio, number of absorbance peaks, peak shape, peak height, etc). The UV absorption spectrum of a Chinese medicinal herb is a complex spectrum formed by the overlap of the characteristic absorption spectrum of multiple constituents. Under certain conditions, the complex combination of the many constituents of a Chinese medicine possesses a certain regular pattern, which shows up in the UV spectrum with a definite specificity and stability. For Chinese medicines that are hard
to distinguish or have close kinship, one can select the multiple-solvent UV spectrum method or derivative spectrum method. The derivative spectrum method is able to eliminate some interference in the original spectrum which is unrelated to the absorbance. It can also resolve the problem of overlapping absorption peaks in the UV spectrum. With the increase in the number of derivatives, the absorption band becomes narrower, the peak shape becomes sharper, resulting in an increase in the spectral resolution.

IR spectrum

Infrared spectrum is an absorption spectrum that reflects the change in the vibrational and rotational energy levels of the intermolecular bondings. IR spectrum possesses high specificity. Every single compound has its characteristic IR spectrum. A Chinese medicine is a mixture of many chemical constituents. Its IR spectrum is the overlap of the absorption in the infrared region (400–4000 cm\(^{-1}\)) by the overall functional groups of each constituent in the mixture. Under the same conditions, the chemical composition in the same Chinese medicine would have a relatively constant quality and quantity, which is distinguishable from other Chinese medicines. Hence the overlapped IR spectrum is stable and specific. Spectrum analysis focuses mainly on the comparison of the characteristic outline (graphical shape of the entire spectrum) of the IR spectrum as the main absorption peaks do not need to be analyzed. Within the scanned wavenumber range, it is sufficient to compare the wavenumbers of the absorption peaks, the peak shape and peak intensity of the same wavenumber, and the differences in the appearance of the “fingerprint region.” With the fast development of instrumental analysis and chemometric and computational techniques, for a complex mixture system like Chinese medicine, many more new methods and new technologies which are even more practical have been developed. Examples are the Fourier transform IR (FTIR), two-dimensional IR correlation spectrometry (2D FTIR), near IR (NIR), Fourier transform–Raman (FT-Raman), attenuated total reflectance FTIR (ATR-FTIR), diffuse reflectance FTIR (DR-FTIR), near-infrared diffuse reflectance spectroscopy
FA Reading

(NIRDRS) and near-infrared diffuse reflectance spectroscopy visualization convolution transform (NIRDRS-CT), which are even more suitable for fingerprint research and construction.

Fluorescence spectrum

Certain components in Chinese medicines, usually molecules with conjugated double bonds or aromatic systems, under UV light, are able to absorb light energy of a certain wavelength, and emit light of an even longer wavelength. When the UV light stops, the photoluminescence instantly disappears. This light is usually luminous visible light, which is called fluorescence. Fluorescence spectrum includes excitation spectrum and emission spectrum. Excitation spectrum refers to the relative efficiency in using different wavelengths of excitation light to cause a substance to emit a single wavelength of fluorescent light. This means that the fluorescent emission monochromator is fixed, while the excitation monochromator scans the wavelength to record the corresponding fluorescence intensity at various wavelengths. Emission spectrum refers to the relative efficiency in using a single wavelength of excitation light to cause a substance to emit a different wavelength of fluorescent light. This means that the excitation monochromator is fixed, while the emission monochromator scans the wavelength of the fluorescent light produced by the substance, to record the fluorescence intensity at various wavelengths. Different molecular structures of a substance absorb UV light and emit fluorescent light of different wavelengths. Its fluorescence spectrum has a definite specificity; the highest excitation wavelength ($\lambda_{ex}$), highest emission wavelength ($\lambda_{em}$) and peak number are characteristic constants used for analysis. Sometimes, to improve the sensitivity and selectivity of the analytical method, weakly fluorescent substances are reacted with some fluorescent reagent to obtain strongly fluorescent products. Despite the small number of fluorescent peaks, one is able to obtain both the excitation and emission spectra of the substance, which provides more information compared to the UV and IR spectra, and is more effective for identification.
Nuclear magnetic resonance

Nuclear magnetic resonance reflects the changes in the energy level of the magnetic atomic nuclei (H atom in H spectra, C atom in C spectra) in a molecule when under the influence of a strong magnetic field and radiofrequency, which produce nuclear magnetic resonance. It is one of the important methods for the identification of an organic compound, providing structural information about the compound, such as each proton’s chemical shift, number and coupling. In Chinese medicine (characteristic whole extraction), under the effect of electromagnetic radiation with a certain frequency, very long wavelength (10 cm to 100 m) and very low energy level, atoms of certain elements (usually H) in the chemical composition are able to absorb electromagnetic radiation to produce $^1$H-NMR spectrum of peak intensity versus absorption frequency. The composition of Chinese medicine is complex. If a certain procedure is used to obtain the total extract of the characteristic chemical component(s), and at the same time, the contents of these characteristic chemical components are relatively constant, then, under the conditions of standardized extraction, a strict correlation would exist between the $^1$H-NMR spectrum of a plant-based Chinese medicine and the plant strain. Some research studies have shown that the $^1$H-NMR fingerprints of Chinese medicines possess a high degree of specificity and reproducibility by carrying out an analysis of the characteristic resonance signal and data ($\delta_{ppm}$) of the $^1$H-NMR.

Mass spectrum

Mass spectrum is a spectrum formed from the line-up of the mass to charge ratio (m/z) of the charged particles (ions). It is the mass spectrum of physical particles that uses a high-speed electron stream to bombard sample molecules, causing them to break down into many ion fragments. Then, in the magnetic field, amid an alternating electric field or vacuum environment, these fast-moving ion fragments are separated. Based on these ion fragments, conclusions are made about the molecular weight and the molecular structure. Under the bom-
barding of electrons, characteristic extracts of Chinese medicines are broken down into fragments in the mass spectrum to produce the EI-MS spectrum of the chemical component in the extract. Since different Chinese medicines contain different components, the molecular ion base peak and the fragment peaks (m/z) shown in the mass spectrum are also not the same. Hence mass spectrometry possesses stronger fingerprint characteristics.

**Chromatography**

Chromatography is the fastest developing and the most widely used modern analytical method in the research area of Chinese medicine. The principle of chromatography is based on the different interphase distribution of substances and the results of their separation. In the recent 100 years of development, various types of chromatographic techniques have been created. Chromatographic techniques have extremely strong separating power and adaptability, and are the top choice in Chinese medicine fingerprinting. The commonly used methods are thin-layer chromatography (TLC), gas chromatography (GC), higher-performance liquid chromatography (HPLC), high-performance capillary electrophoresis (HPCE), high-speed counter-current chromatography (HSCCC), supercritical fluid chromatography (SFC) and circular dichroism (CD).

**Thin-layer chromatography**

Thin-layer chromatography is currently the most widely used method in fingerprinting, as it is fast, economical, reliable, easy to operate, and supports a wide range of applications with good reproducibility. The unique strength of TLC is that it provides a visual image of visible light or fluorescent chromatograms. It can be further combined with a chromatographic scan or digital processing to obtain a chromatogram that shows the different graduation outlines and the corresponding integral data. It is especially suitable for daily analysis and field inspection, and is the main method for Chinese medicine identification. Modern TLC (that is, 2-D chromatography), through the help
of advanced technology and computation technology, has developed to the stage where instrumentation, automation, computation and hyphenation of other chromatographic techniques are used. Special technologies begin to appear, such as high-performance TLC (HPTLC), thin-layer chromatography scan (TLCS), thermal micro application separation (TAS) TLC, thin-layer chromatography–gas chromatography (TLC-GC) and thin-layer chromatography–infrared spectrometry (TLC-IR). Of these, the scan chromatogram obtained by using a fixed wavelength to scan the spots in TLCS is much more objective and accurate than the visual estimation of TLC. On top of that, the peak area of each peak and standard peak ratio can be quantified, which has a greater significance for fingerprint analysis.

Gas chromatography

Gas chromatography allows high resolution and fast analysis. The chromatogram obtained has good reproducibility and good resolution. Since it is a type of closed-system chromatography, there are few influencing external factors, the stability is good, and there is a wide choice of detectors. It is especially suitable for research in Chinese medicine that contains volatile components. In recent years, with the advancement in technology, research and practical needs, new methods and new technologies with even stronger practicalities have appeared. Examples are capillary gas chromatography (CGC), gas chromatography retention index spectrum (GCRIS), pyrolysis–gas chromatography (PGC), head-space gas chromatography (HSGC), gas chromatography–mass spectrometry (GC-MS) and gas chromatography–Fourier transform infrared spectrometry (GC-FTIR). Of these, the most prominent one is pyrolysis-gas chromatography. It is easy to operate, its samples do not require pre-treatment, and it provides a large amount of data. It also has a wide range of applications. For certain Chinese medicines that cannot be directly analyzed with GC, PGC can be used. PGC combines pyrolysis and gas chromatography technology. The sample is put into a pyrolysis device and heated under certain conditions to pyrolyze it into volatile small molecules instantaneously. The molecular products are transferred into the GC
column by a carrier gas. After separation, a reproducible pyrolysis–gas chromatogram which reveals the fingerprint characteristics will be obtained from the recording instrument. GCRIS is based on the GC retention indices of the volatile components in Chinese medicine. From these, the retention indices of certain characteristic components are selected to make up the gas chromatographic retention index spectra, such that the characteristic fingerprint is more distinctive. GC-MS is widely applied and is able to provide “on-line” information about the chemical structure of the main components in the fingerprint. Fast and sensitive, it is a powerful support tool for fingerprinting. The use of supercritical CO$_2$ fluid extraction as pre-treatment for GC-MS has also tremendously broadened the scope for Chinese medicine fingerprint construction.

High-performance liquid chromatography

High-performance liquid chromatography has the characteristics of good separation efficiency, fast speed of analysis, good sensitivity, good stability and reproducibility, wide selection of mobile phase and detectors, and re-usable column, making it extremely suitable for the construction of Chinese medicine fingerprints. HPLC is not limited by the sample volatility and heat stability and is widely used for non-volatile components. With the popularity of HPLC, its applications are increasingly expanding. Currently, it is the main and most commonly used method for the construction of Chinese medicine fingerprints. With the development of new, high-efficiency common detectors, such as the introduction of diode array detectors (DAD), evaporative light-scattering detectors (ELSD), electrospray ionization–mass spectrometry (ESI-MS), MS/MS, the applicability and resolution of HPLC have increased significantly, making possible the analysis of complex samples. From HPLC-DAD, three-dimensional HPLC chromatograms can be obtained. With the peak retention time and spectral data integrated into a single chromatogram, it is suitable for test samples with complex components that absorb light in the UV region. HPLC-ELSD is able to analyze substances that do not absorb UV light. Through HPLC-UV-MS, the HPLC-UV and HPLC-MS
fingerprints can be obtained and the main peaks can be classified. The total ion chromatogram obtained from HPLC-ESI-MS contains even more data, and possesses greater specificity and fingerprint characteristics. HPLC-DAD-MS/MS can be used to establish multi-dimensional fingerprints and total ion chromatograms at the same time.

High-performance capillary electrophoresis

High-performance capillary electrophoresis is a new separation and analytical technique that has been rapidly developing in the last 10 to 20 years. It is a liquid separation technique that uses a high-voltage electrical field as the driving force, the capillary as the separation channel, and separates components of a sample based on their differential mobility and separation behavior. It has the characteristics of high separation efficiency and fast analytical speed. It has many types of separation models, is easy to clean and maintain, and is able to analyze aqueous solutions directly. It is increasingly showing prominent prospects in its applications in the research of Chinese medicine fingerprinting. HPCE is suitable for the analysis of most chemical components, especially the separation of large biomolecules such as peptides and proteins found in animal-based Chinese medicines. The chemical composition of Chinese medicines is complex and diversified; the molecules are of different sizes. HPCE is able to analyze both large molecules and small molecules (such as proteins and phenolic acids) simultaneously. The fingerprint thus obtained has a greater ability to reflect the characteristic components of the Chinese medicine. A variety of separation models are available; the ones that have been developed are capillary zone electrophoresis (CZE), capillary gel electrophoresis (CGE), micellar electrokinetic capillary chromatography (MECC), capillary isotachophoresis (CITP), capillary isoelectric focusing (CIEF) and capillary electrochromatography (CEC). Of these, CZE and MECC are the two most commonly used models in the research of Chinese medicine fingerprinting. The reproducibility of HPCE can be improved through mobility comparison, reference correction and retention index methods.
High-speed counter-current chromatography

High-speed counter-current chromatography is an internationally popular liquid-liquid separation technology. The technology employs the principle of dynamic liquid-liquid separation, making use of two non-miscible solvents moving relative to each other. The two phases which are in dynamic equilibrium will separate the sample components based on their distribution ratio. It has the strengths of high separation efficiency, large overloading capacity, no irreversible adsorption on solid carrier and less solvent consumption. HSCCC technology possesses very good adaptability; since the combinations and ratios of solvent systems are infinite, theoretically it can be used in the separation of samples of any polarity, which is an exceptional strength in the separation of natural products. Its resolution and reproducibility are quite good, and the operation is simple and easy to handle. There is little requirement for sample pre-treatment; good separation effect is achieved with simple extraction or even without pre-treatment. Injections can also be repeated in gradient mode or reversed-phase mode. This technology has a good rate of recovery since HSCCC does not need a solid-state carrier, and therefore the problem of adsorption and degeneration does not exist. So long as the separation conditions are properly adjusted, a good recovery rate can be expected. This is why HSCCC is very suitable for the separation and analysis of Chinese medicine components. Many research studies have shown that HSCCC is able to separate components that are not separable through HPLC. Research on the use of HSCCC in Chinese herbal medicines is still in the initial phase. Currently, it is mainly used in the separation analysis of the active components in Chinese herbal medicines. For the past few years, with the advancement of instruments and methods, the combination of HSCCC and mass spectrometry (HSCCC-MS) has successfully been applied in the separation, analysis and identification of chemical components in Chinese medicine. In the research of Chinese medicine quality control and analysis, especially in the field of fingerprint analysis, HSCCC has broad and favorable application prospects.
Supercritical fluid chromatography

Supercritical fluid chromatography utilizes a supercritical fluid as the mobile phase. Through the use of pressure control, the density of the mobile phase is adjusted to effect the change in the solubility of the substances to be separated, such that the various substances go through dissolution and resolution separately. The supercritical fluid is a powerful solvent, with good fluidity and a fast rate of mass transfer. It combines the strengths of HPLC and GC and overcomes some of their weaknesses. SFC has an even wider scope of application than GC, gives better quantitative results than HPLC, requires a shorter run time and easier sample pre-treatment, and is more readily hyphenated with large analytical instruments, such as mass spectrometry, FTIR and NMR. SFC has a very wide range of compatibility. Theoretically, SFC can be used for any compound that can be separated and analyzed by HPLC. Even for substances that cannot be separated by HPLC, SFC can be used to achieve good separation.

X-ray diffraction

X-ray diffraction is the main method used to study the phases and crystal structures of substances. When a certain substance undergoes diffraction analysis, the substance produces different degrees of diffraction phenomenon, under the effect of X-ray. Its composition, crystal structure, type of molecular bonding and molecular geometry determine the distinctive diffraction pattern of that substance. If the material is a mixture, the diffraction pattern is obtained from the superimposed diffraction effect of each component. As long as the composition of the mixture is constant, the diffraction pattern can be characteristic of the mixture. Though the composition of Chinese medicines (including organic and inorganic multi-phase systems) is complex, when the influencing factors such as the site of cultivation or timing of collection are relatively constant, its composition and hence diffraction pattern will be constant as well. Since the compositions of Chinese medicines are all different, their diffraction patterns have dis-
tinctive characteristics. X-ray diffraction fingerprinting has strong fingerprint characteristics, provides immediate information about the sample composition and gives a stable, reliable pattern. XRD can be classified into the single-crystal X-ray diffraction method and powder X-ray diffraction method. Powder X-ray diffraction is fast and simple, provides a large amount of pattern information, possesses strong fingerprint specificity and multiple evaluation indices (such as diffraction pattern, \(d\) value and relative intensity), and is stable and reliable. It is suited for the fingerprint construction of Chinese medicines. The powder X-ray diffraction–Fourier analysis method is based on converting a diffraction signal through Fourier transformation to obtain the X-ray diffraction pattern and the topology pattern, so as to construct a simpler X-ray diffraction Fourier fingerprint that can reflect the characteristics of the overall structure of a Chinese medicine.

Thermal analysis

During the process of heating or cooling, many substances will undergo physical or chemical changes such as dissolution, solidification, decomposition, reaction, adsorption, desorption and crystal transformation, with the release or absorption of heat. The technology that studies and analyzes such changes is known as thermal analysis. According to the type of analysis, it can be classified into thermogravimetry (TG), differential thermal analysis (DTA) and differential scanning calorimetry (DSC). Of these, DTA is the most commonly used. This method studies the temperature profiles of a sample and a reference material against time or temperature, when both are heated at the same rate under the same conditions. The result of the analysis is shown in a thermogram. Different Chinese medicines will have their own characteristic thermograms.

Electrophoresis

Electrophoresis is the directional movement of charged particles in a buffer under the influence of an electric field. Charged components in
Chinese medicines, such as organic acids, proteins, peptides, amino acids, alkaloids and enzymes, in an electric field of a certain strength and the same period of time, will migrate in different directions (towards the anode or cathode), at different speeds and different distances, according to the nature of the charge (positive or negative), number of charges and the molecular weight of each component. This gives rise to a characteristic band number, band distribution and band colors on the electrophoresis map. Based on the different matrices used, conventional electrophoresis can be classified into paper electrophoresis (PE), polyacrylamide gel electrophoresis (PAGE), cellulose acetate membrane electrophoresis (CAME), etc. Of these, the PAGE method is more commonly used. This method only requires a simple laboratory setup, has high specificity and good sensitivity. High-performance capillary electrophoresis is a tremendous improvement on conventional electrophoresis (refer to section on chromatography). Conventional electrophoresis has an edge in the analysis of Chinese medicine that contains different compositions of peptides and proteins, but its results are more easily influenced by the test conditions. Hence, one should strictly control the consistency of the test conditions.

**Hyphenated chromatographic technology**

Chinese medicine prescription preparations have complex compositions. Very often the use of a single chromatographic method or condition is not able to generate a comprehensive fingerprint, and hence the intrinsic quality of the Chinese medicine is not fully and accurately reflected. Use of hyphenated chromatographic technology generates multi-dimensional and multi-data characteristic fingerprints, which are able to resolve the difficulty of demonstrating the complexity of prescription preparation in its entirety. Multi-dimensionality refers to the link up of several analytical instruments to generate the fingerprints. Each fingerprint provides complementary data, such that a clearer and more complete understanding of the complex sample is formed. Currently the most commonly used method is high-performance liquid chromatography or capillary electrophoresis–diode array...
detector–mass spectrometry–mass spectrometry to obtain multi-dimensional fingerprints. It includes the chromatograms from HPLC or CE (with the retention time of each component), online UV spectrum from diode array detectors, stage 1 mass spectrum (with the mass of each component) and stage 2 mass spectrum (with the characteristic fragment of certain components). Multi-data refers to the characteristic fingerprints which include data in the areas of chemistry and drug efficacy. Chemical data relates to the multi-dimensional fingerprints mentioned above. Drug efficacy data is the relationship between the chemical composition and drug efficacy. Through the determination of the mass ratio of each effective group (chemical groups that are confirmed by drug efficacy experiments) and its active components (each type of compounds), the dose-response relationship of the drug can be deduced from the change in the mass ratio of the chemical components. The construction of multi-dimensional and multi-data characteristic fingerprints not only resolves the difficulty of ensuring the efficacy and quality of Chinese medicine prescription in a more systematic and comprehensive way, it also provides a new solution to the problem of a lack of reference standards for Chinese medicine research. With the rapid development of HPLC-NMR, without the acquisition of pure standards, the identification of the structure of each chemical component in a mixture can be achieved through the use of several hyphenated chromatographic technologies.

**Molecular biology technology**

In recent years, people have been realizing that the nucleotide sequence contains hereditary information required for the building of organic compounds, and for the maintenance and reproduction of life. Different organisms contain different DNA sequences. The same organisms contain the same DNA sequence to maintain the same hereditary traits, and yet possess polymorphism such that every individual of the same species will have many differences in its body. Since the sequence of such nucleotides is relatively stable and different in each individual, modern molecular biology technology can be used to
generate DNA fingerprints. DNA fingerprints are highly specific to the individual. Research methods on the construction of fingerprints using DNA molecules from the polymerase chain reaction (PCR) as markers can be classified into two categories based on the subject of analysis. One is the determination of the polymorphism of DNA genomes, such as random amplification of polymorphic DNA (RAPD) and arbitrary primer PCR (AP-PCR); another is the determination of the polymorphism of DNA fragments, such as restriction fragment length polymorphism (RFLP) and direct DNA sequencing. RAPD and AP-PCR markers are able to determine the polymorphism of DNA when the characteristic DNA sequence is not available, which is more suitable for the construction of DNA fingerprints. RAPD has the advantages of being fast, simple and sensitive. It has been widely applied in recent years and its application prospects are extremely positive. RFLP is the most commonly used method to construct DNA fingerprints. It requires no significant decomposition of the DNA. As this method is time- and effort-consuming, there are limitations to its use. Direct DNA sequencing is, through PCR, the use of a universal primer to amplify a specific fragment of heavily degraded DNA from the medicinal material. The fragments are then sequenced. This method has relatively high repeatability and stability. It is able to minimize errors due to the characteristics of the medicinal material and yet has high specificity.

Each of the methods or technologies mentioned above has its adaptability and specialty in its applications. But there will be some problems when they are used in the analysis of the quality of Chinese medicine. Hence, in the research on the construction of Chinese medicine fingerprints, the choice of method should be based on the actual conditions and the actual analysis. In the case of mixtures, the specificity of UV and IR for identification and resolution is poor; the data processing requires the use of computerized pattern recognition technology or fuzzy mathematics. NMR and MS experiments are very costly; hence their applications are not widely promoted. While the characterization of Chinese medicines using the spectrometric fingerprinting method can be used to differentiate them, unlike chromatographic fingerprinting, it is not able to demonstrate the overall
nature of Chinese medicines, which are complex mixtures with different chemical components of various concentrations. When there are changes in the concentrations of certain components in the complex mixtures of Chinese medicines, due to limitations in sensitivity and selectivity, they are not clearly and accurately reflected in the spectrum. Hence, it is challenging to employ this method in the evaluation of the stability of Chinese medicine quality. While TLC is simple and easy to use, with a high rate of utilization, its main shortcoming is the low “column efficiency” and limited information provided. It is difficult to use TLC to demonstrate a complex system made up of up to hundreds of chemical components. Its sensitivity is also inadequate for distinguishing characteristic fingerprints that are distinguishable only through tiny details. Second, the external environment has a greater influence on an open system such as TLC, hence the standards for external environmental conditions (temperature and moisture) are quite high, and operations should be skilled and standardized. GC is suited for volatile chemical components. HPCE has the advantage of good resolution and is more suited for large molecular compounds, but its reproducibility has to be improved. HPLC fingerprinting is able to adequately resolve each complex component, demonstrate the overall chemical characteristics and express the mass ratio of each component. Multiple data can be obtained from HPLC online analysis. For samples with complex components, with fingerprint characteristics that are distinguishable only through tiny details, it is an effective analytical technique. Hyphenated technology is the most effective method of generating fingerprints, as GC-MS, HPLC-MS, HPLC-MS/MS, etc, are able to provide a large amount of various data, and they fulfill the requirements of analyzing the complex systems of Chinese medicines. However, the instruments are costly and their use is not easy to popularize.

In short, the more prominent or promising methods in Chinese medicine fingerprint research are HPLC, HPCE, PGC, SFC, HSCCC and hyphenated chromatographic technology. Currently, the best method for constructing Chinese medicine fingerprints is HPLC.
Data Processing of Fingerprints

Direct visual analysis and comparison

Based on the fingerprint appearance and the data obtained, direct visual analysis, comparison and evaluation of the characteristics of the test samples and reference fingerprints are carried out. This method is suited for qualitative, rapid identification or is used when there is little data in the fingerprint. Currently, this is still a widely applied method.

Quantifiable data comparison

Quantifiable data comparison refers to the introduction of a relative index, rate of overlap, eight strong peaks, $N$ strong peaks, apparent abundance, etc, which are quantifiable data for the analysis, comparison and evaluation of test samples and reference fingerprints. This method is more objective, accurate and has its practicality.

Chemical pattern recognition technology

Chemical pattern recognition refers to the use of computers to classify or describe a substance based on its chemical composition. The fingerprints of Chinese medicines are relatively complex, and visual comparison is bound to affect the accuracy of the result. If it is combined with statistics, computerized fingerprint analysis and pattern recognition, then its application will be more accurate, reliable, fast and convenient. The main methods include principal component analysis (PCA), simple classification algorithm (SIMCA), nonlinear mapping (NLM), constellation graphing technique, fuzzy information analysis, grey relational grade cluster, hierarchical clustering analysis, fuzzy clustering analysis and artificial neural network (ANN). Chemical pattern recognition technology has been widely utilized in Chinese medicine fingerprint research.
Chemometric methods

Commonly used chemometric methods are quantitative structure-activity relationship, molecular modeling and optimization, and use of resolution ($R_s$), sum of relative peak area (SRPA), WSK which is related to resolution and sum of relative peak area, mutual information and other target parameters that are used as the quantitative evaluation standards for fingerprints.

Computerized software analysis

Based on chemical pattern recognition technology, computerized analysis software, such as the comparison analysis software that comes with infrared spectroscopy, was developed for the ease of study and application. Computer-aided Chinese medicine fingerprint similarity evaluation software has been used in the research and actual work of Chinese medicine fingerprinting.

Fingerprint Setup Procedure

The main steps of fingerprint setup are sample collection, method development, data analysis, sample evaluation and method validation. Sample collection involves collecting sufficient specimens that reflect the sample quality, to ensure that the test sample is representative and uniform. Method development refers to the choice of a suitable method to setup and study the fingerprints. Data analysis refers to the processing of the data obtained, the establishment of the commonalities and differences, and the evaluation criteria. Sample evaluation uses an established method to analyze many more unknown samples over a long period of time, in order to further study the feasibility and practicality of the method.

The Chinese medicinal material fingerprint setup procedure is as shown in Fig. 2.
Fig. 2. Schematic diagram of Chinese medicinal material fingerprint setup procedure.

- Standardize the variant place of origin, time of collection and processing methods for Chinese medicinal material
- Representative Uniform
- Medicinal material samples of more than 10 batches
  - Pre-treatment
  - Test sample
  - Select test conditions
  - Method validation and study
  - Pre-treatment
- Medicinal materials of different variant origins, times of collection and processing
- Untested Chinese medicinal samples
  - Method validation
  - Pharmacodynamics testing
  - Establish fingerprint parameters
- Chinese medicinal material fingerprint
  - Fingerprint analysis
  - Fingerprint data
  - Evaluation criteria
    - Fingerprint test, analysis
    - Quality evaluation
  - Computerized data processing
  - Fingerprint evaluation

Fig. 2. Schematic diagram of Chinese medicinal material fingerprint setup procedure.
1.2

Research Methods and Technical Requirements of HPLC Fingerprinting

Sample Collection

Sample collection is the first and the most essential step in the research of HPLC fingerprinting. Since it is impossible to test all samples of the medicinal material and the growth environment, cultivation conditions, season of collection, processing and storage procedures of the medicinal materials which affect their secondary metabolites (active ingredients), the samples collected must be sufficiently representative. Considering the diversity of biological samples, only when an adequate quantity of a sample is used can we clearly demonstrate its characteristics. Hence, there is a need to collect more than 10 batches of samples.

The meaning of “batches” for drug samples is completely different from its meaning for manufacturing products. It refers to mutually independent samples, that is, samples from the same location, same channel or same collection time should not be considered as separate sample batches, so as to ensure that the test results are representative. Samples from different places of origin or from the same place of origin but of different grades may be viewed as different batches. Samples from the same place of origin may be different if they come from different plantations and land plots. All such samples that come from different sources can be considered as different batches. As the collection of medicinal material samples is limited by subjective and objective conditions, using more samples will lead to greater biostatistical significance.
The basic requirements for sample collection are as follows:

1. First, the medicinal material samples need to go through variant identification. Medicinal materials and their sliced or processed products should conform to national pharmacopeia, Ministry of Health standards, local standards or medicinal material processing regulations. The samples selected should be authentic, good-quality Chinese medicinal materials, or are Chinese medicinal materials produced in places of cultivation that conform to standard agricultural practices.

2. The variant, medicinal part, place of origin, time of collection, preparation and processing methods should be standardized for all medicinal material samples of plant or animal origin. The place of origin and the preparation and processing method should be standardized for mineral medicinal material samples. For Chinese medicinal materials that have multiple variants or multiple medicinal parts, a particular variant or a particular medicinal part should be selected. At the same time, one should carry out the analysis and comparison of the fingerprints of representative samples of the different variants and medicinal parts. This is to determine the common characteristics and specific characteristics of the fingerprints of each variant and each medicinal part. The common characteristics of the fingerprints are used to determine the quality of that particular medicinal material; the specific characteristics are used to determine the type of variant and medicinal part of the medicinal material, and also to determine the quality of that particular variant or medicinal part. Usually, a particular place of origin or a particular season of collection is selected. If multiple places of origin or multiple seasons of collection are chosen, one should inspect the fingerprints of representative samples from different origins or different seasons of collection to look for uniformity.

3. To ensure the authenticity of the samples, there should be a complete and original record of the sample collection, the content of which includes
(a) Name of medicinal material: Medicinal material name and its origin (the scientific name of the original plant or animal, its Chinese name, its Latin name and its medicinal part, or in the case of minerals, its class, family, ore name and main composition).

(b) Sample source: Authentic record of where the sample is from, such as whether it is collected from traditional places of origin or from places of origin with rich resources, or supplied by a GAP base; whether it is purchased from the place of origin, from the market or from agents, etc. This facilitates the purchase of raw materials for manufacturing, and ensures the traceability of test data.

(c) Time of collection and the person-in-charge: Sample collection time and the name of the person-in-charge.

(d) Investigation of goods supply: Information on whether the goods supply is adequate and stable, etc.

(e) Origin identification and the person-in-charge: Identify the variant based on the place of origin and the shape and form of the plants. For a plant of unknown origin, experienced specialists are to identify it from its appearance or microscopic characteristics. For closely related variants or wild variants that are hard to distinguish, one should study the fingerprints to make careful comparisons. If the fingerprints obtained have a high degree of similarity, these variants can be used but the details must be clearly recorded, and when practising GAP in the future, only one type of variant should be specified. If the fingerprints obtained have a poor degree of similarity, then one should identify the variants and use them for cultivation instead. For variant sources with poorly identified commercial products, one should narrow down the scope and gather resources so that only one type of variant is chosen and recorded in the pharmacopeia. In medicine formulation, the “emperor medicine” (君药) and medicinal materials of larger prescription quantities should be the focal point for variant
identification, so as to avoid future difficulties that are unanticipated during the fingerprinting process.

(f) Quality assessment: To facilitate accurate analysis of the chromatograms, and to minimize judgment errors in the test results, the medicinal materials should meet pharmacopoeia or Ministry of Health regulations, and have accompanying detailed records. All medicinal material samples should be serial numbered, and photographs of the medicinal materials should be attached when necessary.

(g) Reserve samples, sample storage and labeling: As the study of HPLC fingerprints is a relatively lengthy process, the samples must be stored in a low-humidity and low-temperature area, away from direct sunlight. Labels should bear serial numbers. The serial number of the collected sample should be identical with the serial number of the storage sample and test sample. The number of reserve samples should not be fewer than three times that of the actual number of test samples, to ensure traceability if there are any discrepancies in the test results.

Preparation of Test Samples

In the preparation of test samples, every step of the process should follow standardized operations. For all the batches, the test sample preparation process should maintain consistency in order to ensure that the sample analysis is precise, accurate and enables reproducibility and comparability of the samples. The main operating procedure and data should be recorded in detail.

Sample Collection

Refer to the 2000 edition of the *Pharmacopoeia of the People’s Republic of China* for the official Chinese method for Chinese medicinal material sample collection to ensure that the samples collected are representative and uniform. The laboratory samples collected should
be consistent with the actual medicinal materials used. For above-ground samples, collect about 0.5–1 kg samples. Weigh the shoots, leaves, flowers and fruit separately and record the approximate ratio. For fruit-based medicinal materials, if the seeds are removed during actual use, the seeds should also be removed from the test samples, and it should be recorded as such. If the apparent mass of the medicinal material is not uniform (such as in the case of dissimilar size or dimension), one should note the representativeness of the sample collected. During subsequent testing, where necessary, one should carry out comparative testing to determine whether there are any significant differences in their compositions, to be used subsequently as a reference for practical applications.

**Weighing of Sample**

To follow conventional requirements, the selected samples must be suitably pulverized, then mixed well. Next, the required amount for testing must be weighed. The usual ratio of weighed sample to selected sample should be 1:10, that is, if 1 g of the test sample is required, it should be weighed out from 10 g of a uniformly mixed selected sample. Since a fingerprint needs to provide quantitative information, the usual requirement for weighing precision is to obtain three significant figures. The method of pulverization, the degree of pulverization and the amount of sample weighed depends on the nature of the medicinal material, the test sample preparation method and other actual conditions.

**Sample Preparation**

This is a crucial step in the sample preparation process. Depending on the physical and chemical nature of the chemical composition of the medicinal material, and the requirements of the test method, one should select an optimized extraction and separation method to prepare the test samples. During sample preparation, one should carry out a study on the different extraction solvents, extraction methods, purification methods, etc, making every effort to retain the chemical
component of the test sample at the maximum limit. This is to ensure that the main chemical component of the Chinese medicinal material shows up in the fingerprints and to reflect the chemical component of the medicinal material in the fingerprints as much as possible. For Chinese medicinal materials from which only one or several components are extracted, extract and separate each chemical component according to their nature and prepare the fingerprints separately.

**Preparation of Solution**

The test sample solution should be prepared by dissolving the sample in a suitable solvent in an apparatus of specified volume, according to specified concentrations (g/mL or mg/mL).

**Storage**

As a normal requirement, the test solution should be freshly prepared as far as possible. If continuous testing is required, the test solution should be stored away from light, and at low temperature, in an air-tight container for a short period, usually not exceeding two weeks. For solutions that are unstable, the storage period should not exceed 48 h.

**Labeling**

Clearly indicate the serial number or batch number on the test sample. It should be the same as the serial number of the collected sample or have a clear correlation to ensure that the data is traceable.

**Preparation of Reference**

To formulate a fingerprint, one should setup a reference material or reference peak, and depending on the nature of the component in the test sample, select a suitable control to be the reference. If there is no suitable control, one may select a suitable internal standard to be the
reference. The reference should be prepared using a suitable method according to the requirements of the test method.

**Choice of Reference**

The reference for the fingerprint is usually selected from more than one main active components or target components, mainly for the purpose of checking the robustness and reproducibility of the fingerprints, and to help in the chromatographic identification. In a situation where there is no clear correlation between the drug and the clinical drug efficacy, the reference helps in identifying and evaluating the characteristics of the HPLC fingerprint. It is not equivalent to the control in quantitative analysis. Fingerprints are usually more complex, hence it is not easy to choose or spike in an internal standard. Fingerprints are not a form of quantitative analysis; hence, the purpose of the internal standard is also not equivalent to that of the internal standard in quantitative analysis. Therefore, it is important to carefully consider whether the use of an internal standard is necessary and feasible. If it is both necessary and feasible, one can consider spiking a suitable internal standard into the chromatogram. The reference should have its name, origin and purity listed. If there is no suitable reference available, one may select stable chromatographic peaks in the fingerprint as the reference peaks. Its chromatographic behavior and related data should be listed, and the chemical structure and the chemical name should be documented as far as possible.

**Preparation of Reference**

Weigh it accurately. Use a suitable solvent and make up the reference solution of specified concentration (g/mL or mg/mL).

**Optimization of HPLC Conditions**

The reason why HPLC has a broad range of applications is that it can be applied to medicinal materials that contain components such as
alkaloids, glucosides, flavanones, organic acids, phenols and lignans. This is mainly because one is able to select the appropriate chromatographic conditions according to the test subject. The main HPLC conditions include the optimal choice of column, mobile phase and detector. One should setup the optimal HPLC conditions, so that the components of the medicinal material test sample can be resolved as much as possible. In other words, the more chromatographic peaks are obtained, the better it is. This will allow the internal characteristics of the Chinese medicinal materials to be fully manifested, providing sufficient information for medicinal material fingerprint evaluation and for its quality assessment.

**Column**

The column may be selected according to the main components of the Chinese medicinal materials. For Chinese medicinal materials that contain mainly alkaloidal compounds, the ion exchange column may be selected. For those that contain mainly polysaccharide compounds, a gel column may be selected. For those that contain mainly steroidal compounds, the C\textsubscript{18} reversed-phase column may be selected. Currently, the column that is most commonly used in medicinal material fingerprint study is the C\textsubscript{18} bonded-phase type of reversed-phase ODS column. The normal-phase column, which is rarely used, is for the separation of homologous or isomeric compounds. The ion exchange column is used for the separation of water-soluble ionic compounds. Some compounds may require the use of an amino column.

**Mobile Phase**

The choice of mobile phase should depend on the actual conditions, and the solvent system with the most optimized method should be used. The best separation conditions should satisfy the following requirements:

1. All the components of the test sample can be analyzed or the sample can be analyzed into as many components (peaks) as possible.
2. All the components in the test sample can be separated satisfactorily.
3. The analysis time is as short as possible.

Optimization of the separation conditions mainly involves optimization of chemical factors such as mobile phase composition, pH value of mobile phase and ion pair reagent concentration. There are many optimization methods; the most representative and relatively more developed methods are the triangle methods (including the Glajch triangle method, tetrahedron method and prism method), graphical methods (including the window diagrams method and overlapping resolution mapping) and direct methods (including the simplex method, complex method and repeated design method). The general consensus is that the triangle methods and the direct methods are suitable for the analysis of both known and unknown samples, while the graphical methods are only suitable for the analysis of known samples.

For the mobile phase, other than optimization of the solvent composition, dynamic changes in the solvent ratio should also be considered. As Chinese drugs contain many complex components, normal isocratic elution is unable to achieve the separation of components of different natures, giving rise to fewer peaks in the chromatogram and insufficient information required for fingerprint evaluation. Under most circumstances, gradient elution should be used. Under suitable gradient conditions, compounds of widely differing natures are all separated. The more chromatographic peaks are obtained, the better it is for the fingerprint evaluation and quality assessment of the medicinal material.

**Detector**

Currently, the detectors available for selection include ultraviolet-visible light (UV-Vis) detector, diode array detector (DAD), fluorescence detector (FD), electrochemical detector (ECD), refractive index detector (RID), evaporative light-scattering detector (ELSD).
Of these, the UV-Vis detector is the most commonly used detector for HPLC. Its degree of sensitivity, precision and linearity range are relatively good. It is suitable for component detection of compounds with π-π conjugated or π-π conjugated structure, while a large number of compounds that are without double bonds are not detected. The DAD is a type of multiple-channel optical system and a new model of UV detector with high resolution. It can simultaneously derive the sample chromatogram and the absorption spectra of every chromatographic component (A-λ curve), to obtain the 3-D spectrum fingerprint chromatograms, and is more suitable for test objects of a relatively complex composition that absorb UV light. This detector is able to evaluate the purity of the HPLC peak, and hence improve the scientific accuracy of the test results. ELSD is able to detect substances that do not absorb UV light, and is suitable for the detection of steroidal saponins and polysaccharide compounds present in Chinese drugs. For those compounds that only exhibit end (lower-wavelength) absorption, such as ginsenoside and astragaloside IV, ELSD is a more suitable detector.

Sample Analysis and Method Validation

Chinese medicinal material samples are prepared using optimized extraction and separation methods, and the analysis should be carried out under the best chromatographic conditions for separation and analysis. For those Chinese medicinal materials that contain the same type of or similar components, one may prepare a single fingerprint. For those that contain complex components, a single fingerprint is not able to completely reflect the intrinsic property of that medicinal material. Through extensive comparison and testing, fingerprints that can sufficiently represent the characteristics of Chinese medicinal materials are obtained, to satisfy the fingerprint requirements of specificity, reproducibility, stability and universal applicability. To confirm the reliability of the test results, the stability of the test sample, the precision of the instrument and the reproducibility of the test method must go through rigorous method validation.
**Stability Test**

This is done mainly to validate the stability of the test sample. Using the same test sample, the sample is tested at different time intervals (0 h, 1 h, 2 h, 4 h, 8 h, 12 h, 24 h, 36 h and 48 h). The uniformity of the relative retention times and the peak area ratios of the chromatographic peaks are studied and the time of analysis determined.

**Precision Test**

This is to validate the precision of the instrument. The same sample is injected at least five times consecutively. The uniformity of the relative retention times and the peak area ratios are studied. In the fingerprint, for each peak that defines the peak area ratio of the common peaks, the relative standard deviation (RSD) of its peak area ratio should not be greater than 3%. The relative retention time of each peak should be within ±1 min of the average retention time.

**Reproducibility Test**

This is to validate the reproducibility of the test method. At least five samples from the same batch number are used. The test samples are prepared according to the selected extraction and separation methods. The samples are tested under selected HPLC conditions. The uniformity of the relative retention times and peak area ratios of the peaks are studied. In the fingerprint, for each peak that defines the peak area ratio of the common peaks, the RSD of its peak area ratio should not be greater than 3%. The relative retention time of each peak should be within ±1 min of the average retention time.

**HPLC Fingerprint Construction**

Based on the number of peaks, peak area (integrated value), peak position (retention time) and other related parameters obtained from the test results of an adequate sample number (more than 10 batches), the common peaks (relative retention time, peak area ratio) are
determined. The characteristic peaks (peak combination) are then selected and the fingerprint constructed. Arabic numerals are used to label the common peaks, and “S” is used to label the reference peak. During the experiment, the 2 h chromatogram should be recorded so as to observe the profile of the chromatographic peaks after 1 h.

The fingerprints of Chinese medicinal materials should be fully representative and specific. Analysis and comparison of representative samples with various places of origin, different grades and specifications or different seasons of collection, etc, should be carried out. From there, the common peak with a relatively constant peak area is identified as the characteristic peak for the Chinese medicinal material. The selected characteristic peaks should possess specificity. For Chinese medicinal materials from more than one source, the specific differences among the variants must be studied.

**Analysis and Evaluation of Fingerprints**

Critical parameters (common peaks, overlapping ratio, N strong peaks, characteristic fingerprint, etc) for analyzing and comparing fingerprints are established according to the information obtained from the fingerprints. The degrees of similarity and difference among the characteristic fingerprints are calculated for the fingerprint evaluation. To develop a feasible standard for quantitative evaluation using HPLC fingerprinting, computer technologies (principal component analysis, clustering analysis, the artificial neural network method or similarity degree analysis, etc) are utilized to analyze fingerprint information and to evaluate the degree of similarity among chromatograms.

**Construction of Important Parameters**

**Common peaks**

Common peaks are identified as those peaks with the same relative retention time in the HPLC chromatograms of different test samples.
Standardization of common peaks

The retention time of the reference is used to calculate the relative retention times of the fingerprint peaks. Common peaks of the Chinese medicinal materials are standardized according to the results of over 10 batches of samples.

Ratios of common peak areas

If the fingerprint of the control is used as the reference fingerprint, the ratios of the common peak areas to the reference peak area are calculated by designating the peak area of the reference as 1. If an internal standard is used as the reference, then one of the common peaks (which should be relatively larger and more stable) is used as the reference peak, and the ratios of the common peak areas to the reference peak area are calculated by designating the peak area of the reference as 1. The ratios of the common peak areas should be relatively stable. The ratio of common peak areas in the test samples of the Chinese medicinal material is compared with the ratio of common peak areas in the standard fingerprint. If a single peak area is more than or equal to 20% of the total peak area, the peak area of that peak in the test sample should not deviate more than ±20% from that in the standard fingerprint. If a single peak area is more than or equal to 10% or less than 20% of the total peak area, the difference should not be more than ±25%. If a single peak area is less than 10% of the total peak area, there is no requirement on its peak area ratio. However, the relative retention times of all common peaks should be standardized. If the common peaks do not achieve baseline separation, the total peak area of this group of peaks should be calculated and used as the peak area. At the same time, the relative retention time of each peak in this group should be standardized.

Areas of non-common peaks

After comparing the chromatograms of the test samples of the Chinese medicinal materials with the standard fingerprint, the peaks with
different relative retention values (those peaks which are not common peaks) are classified as non-common peaks. The total area of non-common peaks should not be more than 10% of the total peak area.

**Overlapping ratio**

The overlapping ratio is twice the total number of common peaks divided by the sum of the number of peaks in the HPLC fingerprint of a Chinese medicinal material and in the chromatogram of the test sample, expressed as a percentage. The formula is

\[
\text{Overlapping ratio} = \frac{\text{Number of common peaks} \times 2}{\text{Number of peaks in the fingerprint + Number of peaks in test sample}} \times 100\%
\]

The overlapping ratio indicates the similarity among the fingerprint chromatograms. The higher the overlapping ratio, the greater the similarity among the fingerprints. Depending on the scenario, a reasonable working range for the overlapping ratio should be specified. The overlapping ratio is an important qualitative parameter which provides reliable evidence for identification of the variants of Chinese medicinal materials.

**N strong peaks**

N strong peaks should be identified based on the actual peaks presented and peak area values that emerged. First, the sizes of the peak areas are listed in descending order and the first N peaks are selected as strong peaks. The total peak area of the strong peaks should be more than 70% of the total peak area. The value of N depends on two factors. The first is the total number of peaks in the chromatogram. The value of N should fall between 1/5 and 1/3 of the total peak numbers. The second is the total peak area of the N strong peaks. In addition, the frequency and the position of the different N strong peaks in the test sample of the Chinese medicinal material should be considered. N strong peaks indicates the relative amount of the main com-
ponents in the Chinese medicinal materials, which provides important information and a basis to assess the quality of Chinese medicinal materials.

**Characteristic fingerprint**

Common peaks are chromatographic peaks that exist in the chromatograms of all the test samples of a Chinese medicinal material, and can be considered a unique characteristic of the materials. That is why they are also called characteristic fingerprint peaks. A characteristic fingerprint is made up of a fixed group of characteristic fingerprint peaks that reflect the multi-component characteristic of a Chinese medicinal material. It provides more detailed information and evidence for the identification of a variety of Chinese medicinal materials, especially for the differentiation of medicinal materials of different species within the same genus or different species that contain the same active ingredients.

The establishment of the characteristic fingerprint requires all the protocols to be standardized. The test samples should be prepared with optimized extracting and separating methods. The tests should be conducted with the best chromatographic conditions for separation and analysis. The characteristic fingerprint peaks are standardized using related parameters from the results of over 10 batches of testing. A series of characteristic fingerprint peaks is selected to construct the characteristic fingerprint of a Chinese medicinal material.

**Degrees of similarity and difference among characteristic fingerprints**

Degrees of similarity and difference are used to quantify the similarity between the chromatograms of the test sample and the standard fingerprint. They combine qualitative information with quantitative data, which can be used as an important basis for final identification.

The degree of similarity between the chromatograms of the test sample and the standard fingerprint is calculated based on the characteristic fingerprint pattern. It indicates the similarity in the fingerprint
patterns of these two chromatograms. The first step is to calculate the degree of similarity of every characteristic fingerprint peak. The formula is

\[
\text{Degree of similarity of a characteristic fingerprint peak} = \frac{A^T_i}{A^F_i} \times 100\%
\]

Then the total degree of similarity of the characteristic fingerprint is obtained by calculating the summation of the degrees of similarity of all the characteristic fingerprint peaks. The formula is

\[
\text{Total degree of similarity of the characteristic fingerprint} = \left( \frac{\sum A^T_i}{\sum A^F_i} \right) \times 100\%
\]

In those formulae,
- \(A^T_i\) is a characteristic fingerprint peak area in the chromatogram of the test sample,
- \(A^F_i\) is a characteristic fingerprint peak area in the standard fingerprint,
- \(n\) is the total number of the characteristic fingerprint peaks, and
- \(\sum\) is the summation symbol.

The total degree of similarity of the characteristic fingerprint should be equal to or less than 1. If the degree of similarity is 1, both of them are completely the same. If the degree of similarity is less than 1, the larger the value, the higher the degree of similarity between them. Any value more than 1 is invalid.

The degree of difference is the difference in peak area between a characteristic fingerprint peak in the test sample and its corresponding characteristic fingerprint peak in the fingerprint chromatogram, divided by the peak area value of the characteristic fingerprint peak in the fingerprint chromatogram.

The total degree of difference of the characteristic fingerprint is obtained by calculating the summation of the degrees of difference of all the characteristic fingerprint peaks. It indicates the degree of differ-
ence between these two chromatograms of the characteristic fingerprints. The formulae are as follows:

\[
\text{Degree of difference of a characteristic fingerprint peak} = \frac{|A_i^F - A_i^T|}{A_i^F}
\]

\[
\text{Total degree of difference of the characteristic fingerprint} = \sum_{n} \left( \frac{|A_i^F - A_i^T|}{A_i^F} \right)
\]

**Computer Analysis of Fingerprints**

**Fingerprint information**

After Chinese medicine samples have gone through sample preparation and chromatographic separation and analysis, the fingerprint data that characterize the chemical components in Chinese medicinal materials are obtained. In the process, data related to characteristic features of a Chinese medicinal material should be collected. The parameters chosen for this purpose can be relative retention time and relative peak area.

For HPLC fingerprints, especially the high-resolution ones, the chromatograms should be prepared from representative samples. Significant characteristic peaks in stable fingerprints should be selected and their peaks should be numbered. Main characteristic peaks should separate well from their neighboring peaks with an effective resolution of over 1.2. Other characteristic peaks should be separated to a certain degree. The peak-to-valley distance should be at least 2/3 of the peak height. If the distance is less than 2/3 of the peak height, the two peaks will be calculated as one peak. To allow the identification and matching of characteristic peaks, the variation of the relative retention time
of a peak should be controlled to be between $\pm 5\%$ and $\pm 10\%$. To study the stability and repeatability of a fingerprint, either a reference should be used or an appropriate peak in the same fingerprint should be selected as the internal standard for calculating the relative retention time and relative peak area ratio ($A_i/A_s$) of characteristic peaks. The relative retention time and relative peak area ratio are used as parameters. If a peak area exceeds 10% of the total peak area, its relative peak area value will be specified. This will be used to establish the reference fingerprint, which will be used to evaluate the authenticity of a sample.

**Analysis of fingerprint information**

HPLC fingerprints of Chinese medicinal materials are very complicated. They contain information which may or may not be useful for classification. Manual analysis may not be accurate. If the method is supplemented with computer analysis and identification techniques, its applications will be expanded significantly. Fingerprint analysis means analyzing the fingerprint, selecting useful characteristic information (i.e., characteristic information which can reflect differences in the chemical components and their concentrations in different samples), and converting it into a digital matrix, which can be recognized by the computer, to obtain the objective criteria for evaluating the quality of Chinese medicinal materials and to provide standards for establishing a Chinese medicine quality evaluation system. Currently, there is no universal standard for fingerprint analysis. An appropriate method for chemical pattern identification should be selected according to the situation. Below are the most advanced and current computer-assisted fingerprint analysis methods.

**Fuzzy information analysis**

This method is used for classification, identification or deduction from fuzzy information. The fuzziness of the information refers to the uncertainty in the classification process, which is caused by the intermediary transition of objective differences. The intermediary transi-
tion of a difference means the process of quantitative change to qualitative change from one side of the change (such as genuine herb) to the other side of the change (such as counterfeit herb). Without a doubt, there is a certain degree of fuzziness in the characteristics of the fingerprints of Chinese medicines. This kind of fuzziness is usually represented by a membership grade, which is used for Chinese medicine authentication. Usually, the characteristics of fuzzy information can be discovered by comparing the definite information and random information. Inexorable laws can be obtained by analyzing definite information, while statistical laws can be obtained by analyzing random information. The characteristic of fuzzy information is that it can provide a fuzzy reference for people to fuzzily identify its corresponding inexorable laws and statistical laws. For authentication of Chinese medicinal materials, the basic procedures for computer-assisted fuzzy identifications are

1. Characteristic information collection: It means collecting digitized characteristic information from the fingerprints of a Chinese medicinal material, which represents its quality. When collecting the characteristic information, the following rules should be followed: Discard, as much as possible, information that will not affect the determination of classification, and keep all the information that will affect the determination of classification. Otherwise, the accuracy of classification will be affected.

2. Follow a certain method to determine the subordinate degree of the characteristics of a Chinese medicine fingerprint for genuine modules and counterfeit modules.

3. Authenticate the Chinese medicinal materials by following the maximum subordinate principle or threshold principle.

Artificial neural network method

This is a brand new information-processing method, which mimics the functions of human brains. Its theory is based on the mathematical model of a neural network. A neural network system contains a lot of basic information-processing units, which are called neurons or ganglia.
There are plenty of complex connections among neurons. The strength of a connection can be represented by a variable weightage. Neurons have learning capacities. Based on certain learning rules, they can change the strength of connections according to problems being processed. A neuron network has its special ability in dealing with problems involving indeterminate relationships, backgrounds and rules. It is extremely suitable for processing the fingerprint information of Chinese medicinal materials.

The back propagation (BP) model is currently the most widely used artificial neural network model. It has several advantages, such as high reliability and high speed of classification. A BP network has multiple connection nodes. It consists of an input layer, a hidden layer and an output layer, which correspond to the input, transition and output of the system, respectively. There are two steps in the BP network’s operation process. First, a training set is formed from a group of samples with known results. For every input of a sample, an output result is obtained according to forward propagation. Second, the difference between the actual output result and expected output result is calculated. The difference will be fed back from the output layer to the input layer to modify the connection weightages of different connection nodes. This procedure is repeated again and again until the differences among every model and every output unit in the whole training set are less than a certain value. For a trained network, the system rule, prediction ability and variable transformation are incorporated in the network in the form of weightages. Processed results will be obtained when information is imported into an input layer of the trained network.

Grey relational grade cluster method

The grey system is a system that contains both known and unknown information. As the fingerprints of Chinese medicinal materials have grey characteristics, they can be processed by the grey relational grade method of the grey system theory. In this method, a new pattern recognition model is established by measuring defined relative correlation. Supposing that there are several Chinese medicinal materials,
since every sample has several quantitative indices of characteristics of fingerprints, the evaluation unit sequence is composed. To use the grey relational grade as the evaluation measurement, the reference sequence should first be selected. Then the relational grade \( r_s \) obtained by comparing it with the optimal reference sequence and the relational grade \( r_t \) obtained by comparing it with the worst reference sequence are calculated. The higher the \( r_s \) the stronger the correlation between the evaluation unit sequence and the optimal reference sequence, indicating a better evaluation unit. The reverse is true for the \( r_t \) value. The lower the \( r_t \), the better the evaluation unit. An ideal evaluation unit should have the highest correlation with the optimal reference sequence and at the same time have the lowest correlation with the worst reference sequence. Thus, the relative relational grade \( r_i \) compared with both the optimal reference sequence and the worst reference sequence is defined as

\[
\frac{r_s}{r_s + r_t}
\]

Obviously, the higher the \( r_i \), the better the evaluation unit. From the value of the relative relational grades of different evaluation units, each evaluation unit can be placed in sequence based on merit. Then the qualities of different Chinese medicinal materials can finally be determined.

Principal component analysis

The principle of the analysis is to factorize the singular values of a fingerprint data matrix into two orthogonal matrices \((U, V)\) and a diagonal matrix \((S)\). The product of the two orthogonal matrices \((U, V)\) multiplied by the diagonal matrix \((S)\) is then calculated. Two of the three values are used as the ordinate and abscissa in the projection diagram. The other value is the projection point in the projection diagram. The changing patterns of the principal components can be identified according to the different places the projection points are located. Samples with a lower total quantity are located to the left end.
of the main component projection diagram. As the position of the sample point is shifted from the left to the right part of the principal component projection diagram, the total peak area in the whole fingerprint becomes bigger and bigger. There is a frequently encountered problem during fingerprint analysis, that is, how to reflect the total amount of extracts of a fingerprint during the evaluation of degrees of similarity. Obviously, it is difficult to reflect the total amount, whether using the correlation coefficient or vector angle cosine. This is due to the fact that when measuring a similarity degree, only the overall similarity of a fingerprint is compared and the differences in the extract quantities are not considered. For example, the chemical pattern recognition method (main component analysis) used in the Evaluation Software for Chinese Medicine (Computer-assisted Similarity Evaluation System) is to project those values into a lower-dimensional space to analyze their nuances, which makes a proper evaluation possible.

Evaluation of a fingerprint's degree of similarity

A fingerprint represents all the characteristics of a sample, which can be used to describe the relationships among samples, visually compare the peak number, peak sequences, peak areas (or peak heights) and relative proportions of the peaks to see if they are similar. All these are just subjective descriptions. The degree of similarity is obtained by measuring the similarity between the fingerprint of a Chinese medicine test sample and the fingerprint of a reference. Comparing the degree of similarity is a kind of fuzzy information analysis method which can quantify the similarity among fingerprints using digitized descriptions that are more objective.

Methods and principles of calculating degree of similarity

Euclidean distance measurement, included angle cosine measurement, correlation coefficient measurement, similarity index measurement and similarity ratio measurement, are often used for calculating the
degrees of similarity of Chinese medicine test samples. Measurements of the correlation coefficient, included angle cosine, similarity index and similarity ratio are used to determine the similarities in the changing patterns of characteristic variables. They are good for providing information on the closeness in relationships among test samples during the authentication of Chinese medicinal materials. The Euclidean distance reflects the closeness in relationships among the test samples that have differences in the characteristic variables. This method is good for the quality control of Chinese medicinal materials.

1. Euclidean distance: Similarity reflects the relationship among investigated subjects and can be measured by distance. The frequently used method is Euclidean distance, or the so-called second-order Minkowski measure.

$$d_{ir} = \sqrt{\sum_{k=1}^{m} (X_{ik} - X_{rk})^2}$$

In this formula, $X_{ik}$ represents the $k$th characteristic variable ($k = 1, 2, ..., m$) of the $i$th sample, $X_{rk}$ represents the $k$th characteristic variable ($k = 1, 2, ..., m$) of the mean vector of a common pattern.

During the calculation of the Euclidean distance, the absolute value of the absolute distance is replaced by a square function, which makes the calculation easier and can emphasize the effects from bigger characteristic variables. The Euclidean distance places extra emphasis on the differences among values of characteristic variables. It does not consider the changing patterns of characteristic variables, i.e. the similarities of those changing patterns.

2. Correlation coefficient: Pearson’s correlation coefficient (i.e. simple correlation coefficient) is used to measure the degree of similarity among test samples in cluster analysis. In fingerprints, the degree of similarity measured by the correlation coefficient is
where $X_{ik}$ represents the $k$th characteristic variable of the $i$th sample, $\bar{X}_i$ represents the mean value of all variables of test samples, $X_{rk}$ represents the $k$th characteristic variable of a common pattern, and $\bar{X}_r$ represents the mean value of all variables of a common pattern.

The correlation coefficient is not related to variable units and is not sensitive to the values of different characteristic variables. The different values of different variables, which have little effect on the changing patterns of characteristic variables, are ignored. The correlation coefficient, also called shape measurement, is to measure the similarities among similar shapes in the changing patterns of characteristic variables of different samples. The degrees of similarity obtained in this way can be used for authenticating Chinese medicinal materials and providing qualitative information.

3. Included angle cosine: Inspired by the similar figures in geometry, the degree of similarity of the included angle cosine between the characteristic vector of a test sample’s fingerprint and vector of the common pattern is calculated based on the multi-dimensional vector included angle. This kind of degree of similarity is related to the changing patterns of characteristic variables in fingerprints, which can provide information for the authentication of Chinese medicinal materials.

$$C_{ir} = \frac{\sum_{k=1}^{m} X_{ik} \cdot X_{rk}}{\sqrt{\left(\sum_{k=1}^{m} X_{ik}^2\right)\left(\sum_{k=1}^{m} X_{rk}^2\right)}}$$

4. Similarity index: Degrees of similarity can be measured by the similarity index when both the mean vector and the standard
deviation vector of a common pattern are considered. This kind of degree of similarity is related to the changing patterns of characteristic variables in fingerprints.

\[ C_{ir} = \frac{1}{m} \sum_{k=1}^{m} e^{-\frac{3(\bar{X}_{ik} - \bar{X}_{rk})^2}{4s_{rk}^2}} \]

5. Similarity ratio: Degrees of similarity measured by the similarity ratio are related to the changing patterns of characteristic variables in fingerprints. The mathematical concept of the similarity ratio is simple and its processing procedures are simple, convenient and fast.

\[ C_{ir} = \frac{\sum_{k=1}^{m} X_{ik} \cdot X_{rk}}{\sum_{k=1}^{m} X_{ik}^2 + \sum_{k=1}^{m} X_{rk}^2 - \sum_{k=1}^{m} X_{ik} \cdot \sum_{k=1}^{m} X_{rk}} \]

Software for evaluating degree of similarity

In current practical work, most degrees of similarity are calculated by computer-aided similarity evaluation software for Chinese medicine fingerprints, which is recommended by the Chinese Pharmacopoeia Commission. There are two types of evaluation software; one is The Fingerprint Analysis System of Chinese Medicine (Professor Cheng Yiyu of Zhejiang University), and the other is Computer-assisted Similarity Evaluation System (Professor Liang Yizeng of Central South University). Both use fuzzy information analysis. Degrees of similarity are calculated using the included angle cosine method. Every fingerprint can be regarded as a group of peak area values or data points at their corresponding retention time. If this group of values are regarded as vectors in a multi-dimensional space, the similarity between two fingerprints is now converted into the similarity between two vectors in a multi-dimensional space. \( \cos \theta \) is used for quantitatively determining and representing the similarity between fingerprints. The closer \( \cos \theta \) is to 1, the more similar the two vectors are.

If there are \( N \) peaks in a fingerprint, it can be expressed in an
$N$-dimensional vector space. If the reference fingerprint is expressed as $X_0 = [X_{01}, X_{02}, \ldots, X_{0N}]$ and $X_{0i}$ is the peak area value of peak $i$, then the fingerprint for testing is expressed as $X = [X_1, X_2, \ldots, X_N]$. The fingerprint of the reference and the fingerprint of the test sample are expressed as two points in the $N$-dimensional vector space. The overall degree of similarity of these two fingerprints is evaluated by calculating their similarity from the included angle cosine of these two points. For example,

$$S(X_0, x) = \frac{\sum_{i=1}^{N} X_{0i} \cdot X_i}{\sqrt{\sum_{i=1}^{N} X_{0i}^2 \sum_{i=1}^{N} X_i^2}}$$

Except for a few species, which should be decided on a case-by-case basis, most of the degrees of similarity of Chinese medicine fingerprints are considered to have met the criteria when the calculated results are within the range of 0.9–1.0 (or expressed as 90–100). If the similarity degree of a Chinese medicinal material is less than 0.9, and when it is difficult to reject the samples just by visual comparison, further evaluation by chemical pattern recognition methods (such as principal component analysis) may be performed.

The detailed protocol for computer-aided similarity evaluation software for Chinese medicine fingerprints is as follows: First, data of statistically significant samples (.txt, .csv file) is imported for establishing common patterns. The data is pre-processed and the resized fingerprints are compared. Peak positions (retention times) are adjusted and peaks are automatically matched. The degrees of similarity of the fingerprints are calculated. After repeatedly deleting and adding samples, the figure and data of a common pattern are obtained and the common pattern template is created. Then test samples (.txt, .csv file) required in the similarity calculation software are generated at the workstation and are processed by the similarity calculation software.

Protocol for Similarity Degree Calculation Software for Traditional Chinese Medicine Fingerprints: Data is imported (.txt file). The data is pre-processed and the resized fingerprints are compared. Peak posi-
tions (retention times) are adjusted and peaks are automatically matched. The degrees of similarity of the fingerprints are calculated. Results are exported and reports are printed.

Protocol for Computer-assisted Similarity Evaluation System: Data is imported. The data is pre-processed, which includes data compression and translation. Peaks are identified and matched. The degrees of similarity are calculated. Principal component analysis can be used to further evaluate different amounts of a certain component.
References


