Method of discovery of effective components in herbals based on evidence by reversed-directed analysis

Conventional methods in this field include identifying and screening compounds in a laboratory setting with cells and animals, and then moving on to clinical trials with promising compounds. To complicate the issue with these methods, effective compounds might not exist in the original herbals, multiple components might be involved with outcomes, long study period and high cost. The method described above might bypass bottlenecks such as, 1) promising compounds are identified from test patients who demonstrate preferred and safe results, 2) more than one compound might be identified, 3) safe and effective compounds might be metabolites and not exist in herbals. The process includes a website to collect clinical information. The promising compounds can be determined by analyzing common components from patients. Both internal and external tests could be done with promising compounds, the entire process would be very efficient and economical.
Title: Method of discovery of effective components in herbals based on evidences by reversed-directed analysis

Author: Jianyi Zhang, MD/PhD
Affiliation: Self-employed physician
Corresponding author:
Jianyi Zhang: 900 S. Washington Street, #112, Falls Church, VA 22046
jianyi.zhang66@gmail.com
Method of Discovery of Effective Components in Herbals Based on Evidences by Reversed-directed Analysis

Introduction

Development and application of chemical compounds to treat illness is a hallmark of modern medicine that is a product of modern chemistry, research and development. Developing a new drug would take an average of 15 years from research to market, and costs about 2 billion US dollars on average. Due to the high cost, only one unknown compound can be tested at a time to prove its safety and efficacy. Study of two or more unknown compounds in the same trial would cost much more, which few pharmaceutical companies can afford.

Traditional Chinese Medicine (TCM) has made an indelible contribution to the health of the Chinese people. TCM is a natural medicine; its effects must have its inherent material basis, which might be achieved by a single or multiple chemical components. Herbals have very complex chemical compositions. Even a single herb consists of hundreds of unknown compounds. TCM is composed of several or even dozens of herbals. Safety and efficacy might be involved with different compounds from various herbals and their metabolites.

The classical method of studying TCM is to isolate numerous chemical constituents in an herbal, then to analyze the activity of the ingredients one by one, in vivo or in vitro. However, this method cannot detect the safety or efficacy of metabolites, nor the synergistic effects among different compounds in TMC. There are numerous methods to analyze chemical compounds in herbals, such as high pressure liquid chromatography (HPLC), liquid chromatography-mass-mass spectrometry (LC-MS-MS). However, these methods can only identify peaks corresponding to chemical compositions. They cannot associate these peaks with effectiveness and safety. Therefore, the material basis is still bottlenecked in a TMC study.

While the method described could rapidly isolate and identify promising, safe and effective compounds from a human being, more importantly, the application of the method would significantly improve new drug development with a much lower cost and reduced time.

Description

To achieve the above goal, a website must be established which is specialized for use by TMC doctors to collect general and clinical information from their patients. Information on such a website should be available for free downloading, which would enable each user to analyze the data. Practitioners would be taught how to use the program in order to conduct statistical analyses.

One mixture may achieve different results in different patients. The mixture could be found safe and effective, or safe, but not effective (or effective, but not safe?). Doctors could, then, analyze the results, adjust compositions, and re-treat patients to improve outcomes (sounds a lot like clinical trials). Practitioners could also compare patients with different forms of treatment to check the effects and possible side effects of herbals. Results could be announced or published, allowing other doctors the opportunity and possibility of repeating these results using the same methods. This proves TMC is a measurable and repeatable treatment option. If preferred results are achieved, the next step is the analysis and discovery of active ingredients. This process is illustrated in FIG. 1.

There are numerous ways of analyzing chemical compositions, such as high pressure liquid chromatography (HPLC), liquid chromatography-mass spectrometry (LC-MS), liquid chromatography mass spectrometry- mass spectrometry (LC-MS-MS), TMC fingerprint, and so on. The human body is composed of thousands of chemical compositions. Using computer programs to analyse blood specimens from individual test subjects, common components among subjects with preferred results could be detected.

Assuming 10 patients (p1 to p10) with the preferred results (shown in FIG. 1), chemical analyses may show that every patient exhibits different components. For example, p1 has 255 compounds, P2 has 302 compounds, etc. Scientists can determine common components using the logic explained in FIG. 2.

Still regarding FIG. 2, if R-Squared ($R^2$) is too small, that means some important components have been lost. Scientists would need to re-analyze the specimens in order to discover which components were lost.

The same common components should then be tested in patients without the preferred responses. If a patient has a common component at a similar concentration as shown in FIG. 1, then no further study would be necessary. A patient without those components or with a low concentration of those common compounds, who did not respond well (as shown in FIG. 1), would require further study.

Statistical analysis could be applied and statistically insignificant compounds could then be removed. $R^2$ could be used to determine how well remaining compounds to explain the efficacy.

Logistical regression could be applied to remove non-significant compositions with binary data as a dependent variable. ANCOVA could be applied for a linear dependent variable, and survival analysis could be used for long-term study. Clinical improvement would be a dependent variable, and the chemical compositions would be independent variables. Non-significant independent variables would be removed if the p-value $>=.05$.

If there are still statistically significant compounds after the analysis, effort would be made to identify these compounds through use of a chemical database. If the structures of those compounds are
known, and are available, they could be used in the next step for in vitro and in vivo studies. If not, those compounds could be purified for further study. That process is explained in FIG. 2.

In those patients who demonstrated no safety issues but with effective results when treated using herbals, the concentration of the identified components could be determined. The next step is to confirm the availability, safety, and efficacy of these compounds.

A chemical database should be searched to check whether identified common compounds are commercially available. If not, the compounds should be synthesized or purified for use in further study.

The safety of these compounds can be determined through clinical trials. The range of these compounds can be discovered by studying blood samples taken at different points in time. Purified or synthetic compounds could be used in animal studies, if they prove safe, and could be used in treatment of patients with the original mixture.

To eliminate inert ingredients in herbals, the synthesized compound could be used in isolation to treat patients. Some of the patients in the last step of FIG. 2 could be treated with components in conjunction with herbals, whereas the rest of the patients could be treated with components only. The purpose of this study is to eliminate non-essential components.

If some patients exhibit preferred results, the compounds would be proven effective and could be used alone clinically, which process is described in FIG 3.

Analyses of the components of each original herbal and mixture of these herbals could be performed, and the results could be compared with other compounds in chemical databases to see if they are structurally known. Active compounds may be present in the original herbals, or the mixed herbals. If the active compounds are not present in these herbals, they might be in vivo metabolites.

**Example 1: Discovery of anti-bacterial components in herbals**

Doctors used herbals to treat patients with E. coli septicemia. Ten patients were treated with the first mixture. Blood collections were done at time points 0, .5, 1, 2, 4, 8, 12, and 24 hours after patients received the mixture. In vitro colony tests were done, and the results were as follows:

At 0 hour, the mean of the colony was 52 with standard deviation (S.D) 11; after 24 hours of the first mixture treatment, the number of the colonies was 222 with S.D. 35. The first mixture was not effective.

Then doctors used the second mixture to treat the same patients. Blood collections were done at the same time points. In vitro tests were done, and the results were as follows:

At 0 hour, the mean was 522 with S.D 41; after 24 hours of the treatment, the mean was 15 with S.D. 3; these results are displayed in FIG. 4. The second mixture was effective; the results are shown in Table 1.
Table 1. Bacteria colonies with two mixture treatments

<table>
<thead>
<tr>
<th>Mixture</th>
<th>Mean at 0 hour (S.D.)</th>
<th>Mean at 24 hour (S.D.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>One</td>
<td>52 (11)</td>
<td>222 (35)</td>
</tr>
<tr>
<td>Two</td>
<td>522 (41)</td>
<td>15 (3)</td>
</tr>
</tbody>
</table>


Based on these results, the second mixture is proven effective and safe. Analysis of blood samples of the patients shows that each patient has many chemical components, but only 4 compounds (C1-C4) are present in all patients.

In an analysis of the blood samples of patients treated with the first mixture, C1 and C2 were found. Chemical databases show that C3-C4 are structurally identified, and commercially available.

The C3-C4 were tested in vitro in a Petri dish and were found to have prominent anti-bacterial properties. The peak concentrations were determined by measuring the blood samples collected at different times from those patients. This allows proper dosage for C3-C4 to be determined. Based on the concentrations, vitro toxicity tests were done with C3-C4 and found without issue. Safety and efficacy tests were done in volunteers. If the patients in the trial showed preferred reactions without apparent side-effects, they could be used to treat patients.

If C3-C4 are non-effective or cause significant adverse effects in the patient, then the chemical databases would need to be rechecked. The structures of C1-C2 are also known and are commercially available. The highest concentration is obtained by measuring blood samples at different times. Toxicity tests in animals were done with C1-C4 and no problems were found. Safety and efficacy tests were performed on volunteers. The preferred result was observed, and no significant side effects were seen. The C1-C4 can treat the disease. This entire process is shown in FIG. 4.

Example 2: Discovery of effective components to treat hypertension with herbals

A TCM doctor treated twenty patients suffering from hypertension with a mixture composed of A, B, C, D, and E herbals. Initially, all patients had high blood pressure, but no other diseases. The test patients came back to the clinic once every three days. After two weeks, blood pressure among the ten test patients returned to normal. There was no change in blood pressure among the ten control patients. Blood collection was done at times 0, 0.5, 1, 2, 4, 8, 12, 24 hours, and 3, 7, 10 days after the treatment.

Logistic stepwise regressions were applied, with findings that compounds A, B, C, and E were statistically significant herbs--$R^2 = .81$. The herbal D was removed from the mixture. This result was confirmed in the treatment of patients by other doctors.

By LC-MS-MS analysis, blood samples from those patients with good results showed hundreds of chemical components in each patient, among which, only four identical ingredients, W, X, Y, and Z were detected. Logistic stepwise regressions were applied, and only ingredients W, Y, and Z were statistically significant--$R^2=0.84$ for these 3 components.
Databases were searched, which led to the conclusion that Y is a known compound and can be synthesized, and Z is an unknown compound. W is a compound contained in herbal A, Y is a compound contained in herbal B, Z is neither present in A, B, C, or E herbals, nor present in the mixture of A, B, C, and E, therefore, Z must be a metabolite of either A, B, C, or E, a vivo metabolite.

The structure of Z was later discovered and can be synthesized. Chemical analysis showed that the highest concentration of W, Y, and Z in each patient is explained in Table 2. All the data in this table are hypothetical.

**Table 2. Chemical analysis shows the highest concentration of W, Y and Z in each patient**

<table>
<thead>
<tr>
<th>Compounds</th>
<th>Mean</th>
<th>Standard deviation</th>
<th>Existed in herbals</th>
<th>Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>W</td>
<td>2.5 pg/ml</td>
<td>.2 pg/ml</td>
<td>A</td>
<td>known</td>
</tr>
<tr>
<td>Y</td>
<td>5.2 pg/ml</td>
<td>.7 pg/ml</td>
<td>B</td>
<td>known</td>
</tr>
<tr>
<td>Z</td>
<td>3.3 pg/ml</td>
<td>0.3 pg/ml</td>
<td>none</td>
<td>Unknown initially</td>
</tr>
</tbody>
</table>

Further analyses are done by the same method in the ten control patients who did not show good results in order to determine if compounds W, Y, and Z were present. Two patients (P1, P2) contained W and Y, with concentrations close to the test patients who displayed marked effects, but no Z compound is present. Three patients (P3, P4, P5) have no W, Y, or Z. The blood samples from patients P6 and P7 have W, Y, and Z, but their concentrations are very low. The other three patients (P8, P9, P10) have W, Y, and Z in their blood, and the concentrations are close to the same as those test patients who displayed good results. The compounds W, Y, and Z were synthesized and used to treat seven patients (P1 to P7) every day, and their blood pressures were checked after 2 weeks.

By biomarker analysis, five types of biomarkers were seen among P1-P10: AAA, BBB, CCC, DDD, and EEE. Type AAA patients only needed Z compound for their treatment, type CCC needed W, Y, and Z compounds for their treatment, while types BBB, DDD, and EEE patients were not treatable with any of these compounds individually or in combination. The results are shown in Table 3:
Table 3. Patients’ responses and their types of biomarkers

<table>
<thead>
<tr>
<th>Patient</th>
<th>W, Y, Z</th>
<th>Add component</th>
<th>Outcome</th>
<th>Possible cause</th>
<th>Biomarker</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1-P2</td>
<td>With W, Y</td>
<td>Z</td>
<td>BP becomes normal</td>
<td>Cannot generate Z</td>
<td>AAA</td>
</tr>
<tr>
<td>P3</td>
<td>No W, Y, Z</td>
<td>W, Y, Z</td>
<td>BP not change</td>
<td>No W, Y, Z receptors or reaction chain</td>
<td>BBB</td>
</tr>
<tr>
<td>P4-P5</td>
<td>No W, Y, Z</td>
<td>W, Y, Z</td>
<td>BP becomes normal</td>
<td>Cannot generate W, Y, Z</td>
<td>CCC</td>
</tr>
<tr>
<td>P6</td>
<td>W, Y, Z very low</td>
<td>W, Y, Z</td>
<td>BP becomes normal</td>
<td>Cannot generate W, Y, Z</td>
<td>CCC</td>
</tr>
<tr>
<td>P7</td>
<td>W, Y, Z very low</td>
<td>W, Y, Z</td>
<td>BP not change</td>
<td>No W, Y, Z receptors or reaction chain</td>
<td>DDD</td>
</tr>
<tr>
<td>P8-P10</td>
<td>W, Y, Z normal</td>
<td></td>
<td>No W, Y, Z receptors or reaction chain</td>
<td>EEE</td>
<td></td>
</tr>
</tbody>
</table>

1. BP: Blood pressure.
2. There was no reason to add W, Y, Z components, since these components existed in those patients within normal range.

The biomarker for ten patients with good results was HHH. In future treatments, these biomarkers should be checked first. If the biomarker is AAA, CCC, or HHH, the compounds W, Y, and Z can be applied for treatment. Patients with biomarkers BBB, DDD, and EEE should not be treated with these compounds as explained in Table 4.

If a patient displays preferred results from herbal treatment, some components must exist for the response. They could be different components which exist in the herbs, or reaction components from different herbals, or metabolic compounds within the human body. Patients showing preferred results should have reaction mechanisms, such as receptors, chain reactions, enzyme subtypes, and so on. Depending on the patient's reaction to herbs, they can be divided into three categories: A, B, and C.

Table 4 below shows the classification of patients based on their response to herbals. Type A and B patients can be treated with these herbals, whereas type C patients cannot be, as these patients lack the necessary reaction chain for the herbals. These types of patients can be identified by scientific methods. All information is hypothetical.
Table 4. Classification of patients based on their responses

<table>
<thead>
<tr>
<th>A: with preferred results in Fig. 1</th>
<th>No</th>
<th>Exist</th>
<th>Exist</th>
<th>XXX, UUU</th>
</tr>
</thead>
<tbody>
<tr>
<td>B: no preferred results in Fig. 1</td>
<td>Yes</td>
<td>good</td>
<td>None</td>
<td>Exist</td>
</tr>
<tr>
<td>C: as B in the above</td>
<td>Yes</td>
<td>Not good</td>
<td>None</td>
<td>None</td>
</tr>
</tbody>
</table>

Note: If the structures of C1-C4 are unknown or they cannot be synthesized, they may be purified first for the studies.

The example cases listed above are not exhaustive, but representative and are given in order to illustrate the application of this patent. There may be other unlimited examples within the scope of the present invention.

Discussion

Conventional methods in this field include identifying and screening compounds in a laboratory setting with cells and animals, and then moving on to clinical trials with promising compounds. To complicate the issue with these methods, effective compounds might not exist in the original herbals. Many original compounds may be modified, metabolized and re-synthesized in living bodies; multiple components might be involved with outcomes. Even though tremendous efforts have been made, few compounds have been discovered and approved by the United States Food and Drug Administration, and made it into the market.

The method described above might bypass bottlenecks such as, 1) promising compounds are identified from test patients who demonstrate preferred and safe results; 2) more than one compound might be identified; 3) safe and effective compounds might be metabolites and not exist in in herbals.

If effective compounds are discovered to exist in nature, which may allow them to bypass approval in the United States by the Food and Drug Administration, they could be patented as a clinical application.

Depending on the nature of the illness, it may only take a few weeks or months to isolate, identify and test promising compounds rather than years. For example, if a group of patients were suspected of having a bacterial or viral infection, and were treated with TCM, the symptoms could improve in a few days. Analysis of blood samples and identification of relevant compounds might take a few days or weeks. If potential compounds are synthesized and available on the market, the further studies would be to verify whether these compounds were still safe and effective in purified forms, the entire process might only take a few months.

During the process, two statistical analyses must be performed. First, remove irrelevant herbals, and second, remove insignificant compounds. These steps would make further studies simpler.
Currently, the failure rate for compounds in clinical trials is around 90%. The high failure rate leads to increased costs in drug development. Much information about the method comes directly from the patients, such as, concentrations of relevant chemicals, profile of side-effects, and pharmacokinetics. This information could aid in reducing, or even eliminating some pre-clinical and clinical studies.

Successful isolation and identification of effective compounds would make industrial production of these compounds possible, which, in turn, means the quality and quantity of these compounds could be well-controlled.

This method provides a complete new way for new drug development. Application of this method will tremendously impact the pharmaceutical industry.
1. Set-up an Electronic Health Record specifically for TMC providers.

2. Providers input initial information.

3. Statistically analyze the data, and find out herbal patterns which are safe and possibly effective, divide patients into ones with preferred response & ones without.

4. Based on the obtained information, providers adjust his management to improve outcomes, and re-cycle the process.

5. Preferred outcome could be available for other providers to see results repeatable.

Yes

Confirm the results?

No

Further analyze relevant effective components in next step.

Figure 1. Digitize TMC which provides information for safety and efficacy of herbals

P5 in preferred response in figure 1 has 198 chemical compounds, only compounds identical with those 198 compounds kept in these other patients.


P3 has 138 chemical compounds, only compounds identical with those 138 compounds kept in these other patients.


Repeat the process above.

After the 10th selection, all patients have same common compounds.

Is there any compound statistically significant?

R² too small? Yes

Go back the step 1 in figure 2, re-separate and analyze compounds.

No

Check the patients in the figure 1 with poor results and find out patients with same compounds at similar concentration.

Remove these patients, treat the rest of the patients (see 5a and 5b in figure 3).

Figure 2. Selection common compounds among patients with preferred outcome.
1. Find out common compounds among those patients (see figure 2).

2. Check database to see if they are known?
   - Yes: Whether they can be synthesized?
     - Yes: 4. Synthesize these compounds.
     - No: Go back point 1. in the figure 1.
   - No: 3. Purify those compounds.

3. Purify those compounds.
   - No: Go back point 1. in the figure 1.
   - Yes: 4. Synthesize these compounds.

4. Synthesize these compounds.
   - Yes: Some effective, some not ineffectively
     - Some effective, some not ineffectively: Further study differences between patients, see figure 7 and 8.
     - Completely ineffective: Go back point 1. in the figure 1.
   - No: 5a. Treat patients with those compounds together with the herbals.\(^1\)

5a. Treat patients with those compounds together with the herbals.\(^1\)

5b. Treat patients with those compounds together without the herbals.\(^2\)

---

Figure 3. Treatment of patients with identified compounds
1. A group of patients has E. Coli septicemia.

2. Treat the patients with the first herbal mixture. Collecting blood at time 0, 0.5, 1, 2, 4, 8, 12, 24 hours of the treatment. Culture blood collected at 0 and 24 hours, and found out the mixture ineffective.

3. Treat the patients with the second mixture. Collecting blood at time 0, 0.5, 1, 2, 4, 8, 12, 24 hours of the treatment. Culture blood collected at 0 and 24 hours, and found out the mixture effective.

4. By analysis, there are 4 common compounds (C1-C4) discovered in these patients.

5. By analysis, there are 2 same compounds (C1-C2) found in these patients treated in the first mixture, C3-C4 are new components after treatment with the second mixture. Check database, and know that C3-C4 can be synthesized and available.

6. Test C3-C4 in vitro to see if they are effective?

7. Test C3-C4 in volunteers to see if they are safe?

8. Based on dosage above, treat patients with C3-C4 & collect blood at time 0 and 24 hours to see they are effective?

9. Check database to see if C1-C2 are available?

10. Test C1-C4 in vitro to see if they are effective?

11. Test C1-C4 in volunteers to see if they are safe?

12. Based on dosage above, treat patients with C1-C4 & collect blood at time 0 and 24 hours to see they are effective?

C3-C4 can be used to treat patients.

C1-C4 can be used to treat patients.

Try to synthesize C1-C2, then go to step 10.

Go to step 10.

Go back step 4, find out more effective compounds.

Give up C1-C4 or find out compounds for safety.

Go to step 4, find out more effective components.

Figure 4. Discovery of effective anti-bacterial compounds