TOWARDS A MODEL FOR PLANNING CLINICAL RESEARCH IN ORIENTAL MEDICINE

Kell N. Julliard, MA,1,* Claudia Citkovitz, MS, LAc,2 and Douglas McDaniel, MTOM, LAc3

Background: Oriental medicine (OM) is widely practiced internationally and embraces many schools of thought. Western medical research is currently struggling to understand OM in purely biomedical terms, with limited success.

Objective: We propose a research model for applying Western research methodologies to OM in a way that respects its theory and modes of clinical application. This would facilitate systematic investigations of OM’s specific assumptions and make explicit the way OM studies could build on each other.

Methods: To develop this model, the authors extracted key assumptions of Western research methodology germane to clinical research, put them in a developmentally logical sequence, and related them to the diagnostic and clinical processes of OM.

Results: The model categorizes studies into seven levels. Foundation studies (level one) establish the conceptual basis for OM research by establishing the internal validity of its basic “truth statements.” Measurement studies (level two) determine how OM identifies and measures diagnostic indicators, treatment outcomes, and other basic aspects of health. Group studies (level three) describe populations in ways meaningful to their health. Pattern/diagnosis studies (level four) identify and define OM patterns of disharmony. Treatment technique studies (level five) describe particular techniques or principles of treatment, their indications, and rationale. Treatment effectiveness studies (level six) evaluate techniques of treatment, often by comparing the results of one technique with those of another in similar patients. Systematic reviews (level seven) draw together studies on the same topic to see if conclusions are thereby strengthened.

Conclusion: The levels can be used to establish relationships between already published studies, determine if sufficient background research has been done to enable a study idea to be carried out, and generate ideas for future studies.

Key words: Research methods, research design, investigative techniques, complementary therapies, acupuncture therapy, Oriental Traditional Medicine, Chinese Traditional Medicine

INTRODUCTION

In contemporary biomedicine, research often informs clinical practice; new drugs, diagnostic methods, and medical devices are often tested experimentally before they are applied in the clinic. In the case of Oriental medicine (OM), however, American practitioners legally perform techniques—such as acupuncture, moxibustion, Chinese herbology, tuina, gua sha, Chinese dietary therapy, Qigong, and cupping—that have been stable elements of Chinese clinical study and practice for over 2,000 years but have not been thoroughly validated by biomedical research methods. This remains so in spite of the exponential increase over the last decade in the amount of OM research being conducted.

This lack of research validation of OM is troubling in light of research’s influence on acceptance of new techniques by the biomedical community. Oriental Medicine appears to be successfully surviving its transplantation to the West without research validation, as evidenced by increasing use by the public and a burgeoning number of practitioner training programs. However, continued growth of OM practice without a corresponding increase in biomedical understanding and acceptance would be undesirable and potentially dangerous, as patients underreport alternative medicine usage to biomedical practitioners who are not perceived as understanding or supporting it.1

As emphasized by Walach et al,2 when the treatments under consideration have been in use for some time, as is the case with OM, evaluation of treatment effectiveness starts with observational and other types of nonexperimental research designs. Most current research simply attempts to evaluate OM in biomedical terms, whether its concepts are able to be studied within the biomedical framework or not.

The assumption that standard biomedical research methods are adequate to address complementary and alternative medicine (CAM) and OM questions lacks justification. Conducting CAM research presents challenges not easily met by conventional study designs.3,4 Research designed to establish causal links, such as randomized controlled trials, might not lead us to the best CAM research evidence.5,6 This is a particular danger for OM, which comprises a vast international body of clinical and theoretical work, only a fraction of which is represented in the acupuncture techniques typically studied by randomized

1 Lutheran Medical Center, Brooklyn, New York; Department of Family Practice, State University of New York Downstate Medical Center, Brooklyn, New York
2 Department of Acupuncture and Oriental Medicine, Touro College, New York, New York; and Lutheran Medical Center, Brooklyn, New York
3 New York Chiropractic College, Department of Acupuncture and Oriental Medicine, Seneca Falls, NY; Pacific College of Oriental Medicine, New York, New York
* Corresponding Author. Address:
Lutheran Medical Center, 150 55th Street, Station 2-30, Brooklyn, NY 11220
e-mail: kjulliard@lmcmc.com
controlled trials (RCTs). Indeed, by imposing the culturally specific intellectual constructs of biomedicine onto OM research, we may be obscuring exactly those areas where it could benefit us most. New research strategies and a variety of research methodologies (especially those that take whole systems into consideration) are needed to explore the relationship between orthodox and unorthodox medicines.7-9

In mentoring acupuncture masters students in their thesis research, the first author (K.N.J.) observed that many of these students wished to carry out randomized controlled trials to prove the effectiveness of OM in treating many conditions. Not only did they have little awareness of the existence and value of the many other research methodologies, they did not understand how to assess whether sufficient foundational research had been done to make appropriate design of a clinical trial possible for a given condition. They also had great difficulty adapting the theory of OM to fit biomedical research models. This situation prompted the first author to develop a conceptual tool that would assist these masters students and potentially other researchers in successfully addressing these issues.

Of course, the research literature has many helpful and well-considered resources that offer methodological advice for CAM research in general and OM research in particular. For instance, Jonas’ concept of an evidence house arranged the main research domains of scientific investigation in a way useful to both complementary and biomedical systems.5 Its semihierarchical but flexible approach offered research domains (“rooms” in the “house”) for investigating the relevance and utility of healthcare practices, as well as domains that seek out causal attributions and mechanisms of action. Building on this evidence house model, Lewith et al9 added concepts related to model validity of research in CAM (ie, how accurately a study takes into account the assumptions of the CAM system under investigation). These authors also explored how key conceptual issues can differ radically between biomedical and CAM systems, such as their dramatically different approaches to diagnostic classifications.

Also addressing CAM in general, Thomas and Fitter1 presented a framework that outlined the important types of research questions and linked them to research designs of use in answering them. The framework of Thomas and Fitter had three main divisions: (1) proving studies that address the efficacy, effectiveness, and safety of interventions; (2) exploratory studies that lay the groundwork for conducting proving studies; and (3) improving studies that aim to improve practice through enhancing understanding of the clinical process. Similarly, the five-data box model of Wittmann and Walach10 provided a way of categorizing research designs for CAM that spanned observational, experimental treatment, nonexperimental treatment, and outcome and evaluation research.

These approaches have much to offer research in OM through their flexibility, practicality, and wide perspective on the way healthcare practices are implemented. Even so, their intention was not to provide a systematic way of taking the unique theories of OM into account while utilizing the strengths of Western research methodologies. The work of MacPherson et al11,12 helped remedy this by highlighting how biomedical methodologies such as case-control and cohort studies and randomized controlled trials can be used to investigate issues important to traditional East Asian systems of medicine. Also aware of this need, Sherman13 in 2004 took a fresh and invigorating approach by identifying top research priorities for the acupuncture community. Because research methods follow from the questions that are asked, the pithy questions posed by Sherman should give rise to new or modified research methods in the attempt to answer them.

Focusing on acupuncture specifically, Lao et al14 in 2001 offered a systematic, stepwise approach to investigating efficacy. These authors recommended adopting the phased approach of the U.S. Food and Drug Administration to develop clinical trials of Chinese medicine. Along with providing valuable suggestions for addressing some of the thorny issues of Chinese medicine research, they identified a need for an integrated research model of Chinese and biomedicine and recommended research within the medical paradigm of Chinese medicine itself. In posing such questions as How valid are the diagnostic methods of traditional acupuncture? and What does each part of the acupuncture consultation contribute to the overall benefit?—White,6 like Lao et al,14—strongly suggested that acupuncture be researched on its own terms. The focus of White’s article, however, was not on this issue but rather on how to best conduct explanatory (tightly controlled, causal) clinical trials from a more biomedical point of view.

Similarly, an article on developing a research strategy for acupuncture by Lewith et al15 astutely addressed the need to look at acupuncture as a whole system, taking into account nonspecific, placebo effects, the difficulty of measuring the effect size of acupuncture, and the importance of looking at effectiveness as well as efficacy. Like Bell et al,4 Campbell et al,16 and Verhoef et al,17 these authors also recommended using qualitative research within randomized controlled trials to better understand patient-perceived benefits of treatment.

Thus, although many research models related to CAM and OM have been developed and much good advice for linking OM research questions with research methods has been offered, a model is still needed that can harness the powerful biomedical tools of research design without importing assumptions that may be invalid for OM.4 Through such a method, OM could be researched on its own terms, “within its own medical paradigm” and using its own rich intellectual resources, but in a way still recognizable to biomedical trained readers as evidence based.14 This paper explores the conceptual foundations for such a model as applied to clinical research.

The seven categories, or levels, proposed here build from fundamental, often tacit assumptions about health and disease (level one), through techniques of measurement (level two), description of groups (level three), diagnosis and pattern identification (level four), description of treatment techniques (level five), evaluation of treatment effects (level six), up to the systematic reviews that are often considered the highest level of evidence (level seven). If applied by OM researchers, this framework could serve several purposes. First, this framework could be used to investigate specific assumptions and observations from various types of OM systematically. Second, this framework makes explicit a way that OM studies could build on each other. Third, it could improve our ability to design and critique research studies by giving us a concrete way to determine whether
the foundation has been laid for answering a particular research question. Finally, although this framework allows for the integration and interpenetration of OM and biomedicine, it also provides a structure for researching OM purely on its own terms.

To develop this system of seven levels, the authors attempted to extract key assumptions of Western research methodology germane to clinical research, put them in a developmentally logical sequence, and relate them to the diagnostic and clinical processes of OM. Thus, issues of measurement, which figure prominently as a prerequisite for almost all biomedical research, appear early in the levels, whereas issues related to treatment, which depends on the ability to measure, assess, and diagnose as well as on the description of techniques of treatment, appear later. The purpose of this paper is to describe the levels and their framework in a way friendly to OM, CAM, and biomedical students, faculty, and practitioners, so that they can use it productively in their own research and critique of the clinical research literature.

LEVEL ONE: FOUNDATION STUDIES
Foundation studies establish the conceptual and philosophical basis for research, the assumptions upon which all other studies rest. Foundation studies can explore such topics as the elements of OM that most need to be measured, the nature of validity within OM, or the use of theory to generate and prioritize specific OM research questions.

Each system of medicine evolves in a particular society with its own social, cultural, and intellectual history. The scope of medicine’s inquiry and the shape of its daily practice are determined largely by cultural constructs—that is, assumptions about what constitutes reality and what does not. In practice, these assumptions function as conceptual inclusion/exclusion criteria, applying tacitly to every experiment run for that system of medicine. Foundation studies seek to render these usually tacit assumptions explicit, thus making epistemology a legitimate subject of inquiry rather than a silent gatekeeper.

Foundation studies in contemporary literature commonly explore issues such as OM’s concept of qi and meridians and its holistically based analysis of health and disease. Although these issues are important as discrete areas of investigation, more fundamental issues exist as well. For instance, biomedicine strives to perfect a single internally consistent model of medical causality, whereas OM utilizes multiple alternative models that can coexist within a single patient-practitioner relationship. Just as Chinese cultural and civic life was built on a pastiche of ideas and symbols drawn ecumenically from Buddhist, Taoist, and Confucian sources, OM practitioners select diagnostic and treatment techniques from various schools of thought widely separated in place and time, without undue concern about theoretical contradictions. The acceptance in OM of multiple coexisting models needs to be taken into account in developing research approaches. Because this standpoint allows the appropriation of technical strong points from a given system without being limited by apparent inconsistencies with other systems in simultaneous use, it is arguably the best standpoint from which to usefully integrate OM and biomedical practice and is used throughout this paper. Rather than being dismissed as primitive or imprecise, this approach should be understood as a significant piece of epistemological equipment that could also prove useful to biomedicine in any future attempt to integrate post-Einsteinian understandings of the physical world into its current, predominately Newtonian models. Foundation level studies address such conceptual and epistemological issues to create workable and valid theoretical bases for research.

For guidance in conducting foundation studies, see chapters 1 and 2 of Portney and Watkins, and articles by Caplan, Ballentine, Engel, Tonelli and Callahan, and Jonas. Cassidy gives examples of how qualitative research can be used to illuminate foundational issues in OM.

Sample Research Ideas
(1) What concepts in OM theory are most difficult to understand for biomedical researchers involved in acupuncture research, and (2) how can they be successfully modeled in contemporary terms? (3) A fundamental concept in OM is the relationship of acute and chronic illness to inclement weather. What interventional or observational methods could be used to evaluate the clinical validity of this concept?

LEVEL TWO: MEASUREMENT STUDIES
Measurement studies determine valid and reliable ways for OM to identify and measure diagnostic indicators, treatment outcomes, and other basic aspects of health. Measurement is the front line of medicine. It is where abstract concepts of health and disease (as defined by level one studies) meet the infinitely varied specifics of human life and suffering. In this encounter, the nature and amount of data collected defines a system of medicine as clearly as any subsequent analysis or treatment. Measurement studies focus on the way this data is collected and interpreted—the validity of the way of measuring (Am I measuring what I believe I am measuring?) and its reliability (Will I get the same result if I measure the same thing more than once?).

One of the most conspicuous achievements of biomedicine is the extraordinary degree to which it uses technology to measure things beyond the reach of human senses—from radiographs and magnetic resonance imaging to assays and cell cultures. Oriental Medicine, on the other hand, has generally sought to refine the senses themselves by training practitioners to use them attentively and systematically. The traditional four examinations consist of looking (evaluating characteristics visible to the eye), listening and smelling (evaluating such characteristics as voice, respiration, and bodily odor), asking (evaluating through questioning), and touching (evaluating through such actions as palpation and taking the pulse). In these examinations, observations are made in a methodical way, so that over years of training and practice, practitioners’ sensory observations (measurements) of such general and apparently subjective parameters as pulse quality, smell, tongue and face color, flesh tone, vocal timbre, and the presence of spirit (shen) in the eyes become increasingly acute and repeatable. The validity and reliability of the ways those distinctions are made in the four examinations by the same and different practitioners are beginning to be seriously questioned in the research literature, indicating a great need for solid measurement studies.
Because OM is a form of medicine, the assumption is often made that its ideal measurement model is that of biomedicine. Maintaining model validity (congruence of the forms of measurement with the theoretical model of the system), however, is important in all aspects of CAM research, including ways of measuring. Because OM is holistic and has psychological and behavioral aspects, the measurement models of psychology and the social sciences may also be useful. Under these circumstances, qi or yin could be considered as constructs, which are abstract or general categories inferred from specific circumstances. A measurement tool for the various states of yin or qi could be developed much like those used to measure the constructs of anxiety or assertiveness or introversion. One of the values of such tools is that they can combine items that are considered physical, emotional, mental, social, and spiritual from a Western perspective. To learn more about constructs, consult such works as those by Wittman and Walach, Edwards and Bagozzi, and Kelly. Along with the methods of measurement used in psychology and sociology, qualitative methods of measurement can also be used in innovative ways, both on their own and in conjunction with quantitative research.

Resources for carrying out measurement studies and establishing reliability and validity include those by Portney and Watkins, Miles and Huberman, Morse and Field, Kapchuk, Verhoef et al, and Wittman and Walach.

Sample Research Ideas

1. Would controlled lighting and a calibrated color chart result in better interpractitioner agreement in measuring tongue color?
2. Does the measurement of shen disturbance differ between practitioner-patient pairs who share the same ethnicity versus those who do not?

LEVEL THREE: GROUP STUDIES

Group studies describe populations in ways meaningful to their health. A population is a group of people somehow logically related, such as persons who live in a certain city or region or who come from the same ethnic group, patients at a particular treatment center, or persons sharing a particular trait, pattern, or diagnosis important to OM (eg, crimson tongue, choppy pulse, or yin deficiency). Group studies often set out to pin down how frequently something occurs in a population or describe the specific characteristics of a population that need to be better understood. Of course, the specific ways of describing those characteristics need to have been established as reliable and valid through measurement studies (level two) for the level three studies to be undertaken with confidence.

Philosophically, OM is chiefly defined by the interplay between two systems of reasoning, the yin and yang system and the five phases system (which is based on the five processes of growth/wood, reaching the maximum/fire, balance or neutrality/earth, decline/metal, and maximal rest preparing for change/water). Both systems function by dividing the known universe into qualitatively distinct groups for the purpose of systematically analyzing change. Thus, the most fundamental way of describing groups is in terms of their balance of yin and yang or the particular disturbances or emphases of one phase compared with another.

Clinically, however, any element assessed during the four examinations can be used to group people (healthy or sick) according to their constitutional tendencies or transient symptomatology—for instance, sensitivity to heat or cold or having a slippery pulse or a tongue with a dark tinge. Constitutional types are particularly important in OM, as pathology is seen to express itself primarily as an exaggeration of existing constitutional trends. For example, individuals with marked constitutional tendencies to heat or cold can be expected to respond with markedly different symptoms when infected with the same rhinovirus.

Group studies abound in biomedicine, the Framingham heart study being a classic example. Their important purposes—to better understand risk, prevention, and prognosis—could just as easily serve OM. Observations of groups abound in the classic OM literature. For instance, those who have recently been out in cold weather would be more prone to chills, mild fever, headaches, and body aches, and those whose sexual behavior is excessive would be more likely to complain of lumbago, dizziness, or diminished vitality. Group studies similar to biomedical cohort studies could help us understand the functioning of risk and prevention in OM in a whole new way.

Resources that are helpful in carrying out group studies include Wittman and Walach, especially with respect to cohort studies, Portney and Watkins, and epidemiology texts such as that by Friis and Sellers.

Sample Research Idea

Oriental Medicine would predict a prevalence of physical signs such as distended sublingual veins and a “wiry” pulse in groups reporting emotional problems such as frustration, stress, and hating one’s job. The determination of this correlation through research could shed light on the use of OM as preventive medicine.

LEVEL FOUR: PATTERN STUDIES

Pattern (diagnosis) studies identify and define OM patterns. In OM, practitioners identify which of the many potential patterns of disharmony apply to a particular patient by assessing physiological and psychological attributes of that individual. These patterns are derived from the combination of the patient’s signs, symptoms, and general characteristics and are considered to be manifestations of an underlying imbalance of some kind. They are described using the elements of OM, such as yang and yin, repletion and vacuity, channels, organs such as stomach and lung, and fundamental substances such as qi or Blood. Diagnosis (here taken to mean the identification of a pattern of disharmony causing distress) represents the practical culmination of level one (foundation studies), level two (measurement studies), and level three (group studies). In the process of OM diagnosis, human intelligence, careful observation, and experience are used to draw a relationship between membership in a
particular diagnostic group and the symptoms suffered. Pattern studies ultimately must rely on group studies (level three) because they establish similarities and differences between groups, the most fundamental groups being those who share a pattern or constitutional type with those who do not.

Many differences exist between the diagnostic processes of OM and biomedicine. Among the most fundamental is that biomedical diagnosis attempts to isolate a definitive physical cause for a particular disease, whereas OM in many cases points to a constellation of related, self-perpetuating functional disharmonies in the system of the individual. Another important difference is that biology strives always to diagnose from a single conceptual model—revising itself to reflect the most current scientific understanding available—whereas OM adds new models (including biomedicine) to itself without discarding the old. The result is a certain relativistic pragmatism in diagnosis; in OM, differently trained expert practitioners may well diagnose the same patient differently. The only objective standard by which their claims to correctness can be evaluated is the success or failure of treatment based on the diagnosis.

To further complicate matters, the diagnostic and treatment processes in OM frequently overlap each other, as they sometimes do in biomedicine (as for example with esophageal reflux as a cause of chronic cough discussed by Glassziou). In OM, the pulse is one indicator of the effects of treatment, so acupuncturists frequently check the pulse between needle insertions to verify that their treatment is working. They may in fact revise the diagnosis based on which treatment is successful and how indicators such as pulse respond to treatment. Furthermore, in systems such as Japanese acupuncture, practitioners may use such measures as pulse and palpation of the abdomen to identify and resolve areas of stagnation, effecting systemic and symptomatic changes, without ever performing any single action discretely identifiable as diagnosis. Thus, even the applicability and usefulness of diagnosis and pattern identification are fertile areas for research.

Some resources related to conducting pattern studies are as follows: Schnyer et al. for ideas about how the Delphi technique can be applied to the study of patterns and assessment in general; resources related to the development of assessment instruments such as the introduction of the Diagnostic and Statistical Manual of Mental Disorders (American Psychiatric Association), global assessments such as the SF-12, and texts on the development of assessments such as those by Thorndike and Hopkins.

Sample Research Idea
How do the indicators for liver qi stagnation differ among expert practitioners diagnosing the same panel of patients?

LEVEL FIVE: TREATMENT TECHNIQUE STUDIES
Treatment technique studies describe in detail a particular technique or principle of treatment, its indications (patterns, diagnoses), and its rationale. Once a clear means of assessing patterns exists, treatment related to those patterns can be evaluated. Of course, many treatment techniques are found in standard texts, which can be referenced for most applications of OM. Technique studies are particularly needed when the technique is new, has a new application, or is underutilized or misunderstood.

Biomedical treatment technique studies often include the following: a rationale for the technique in treating a patient with a particular symptom, pattern, or diagnosis; a statement about why the technique needs to be published at this time; the theory underlying the technique; a clear description of the technique; and a case report or case series illustrating the characteristics of typical patients, details of utilization, and the outcomes (good and otherwise) of the technique.

The classical OM literature in fact consists largely of advice regarding treatment techniques, transmitted generation by generation to the present in technical manuals and a large volume of case studies. Western studies of OM have tended not to validate this advice on its own terms, instead looking to randomized controlled trials of OM protocols for biomedically defined disorders (which seldom sufficiently consider the applicability or effectiveness of the technique for specific individuals). Furthermore, techniques of treatment are often inadequately referenced in studies relying upon them. An important methodological advance in this regard has been the treatment manual, as described by Schnyer et al. Using this approach, state-of-the-art manuals valid from an OM perspective can be developed for specific forms of treatment. The creation and validation of such manuals is one important subset of level five studies. Another is the in-depth case study, which should be reinstated as a primary learning tool for OM students and practitioners.

One resource for writing technique studies is by Stux and Birch, which provides advice regarding components of acupuncture techniques that need to be reported. The surgical and osteopathic literature offers many examples of technique studies. Qualitative methodology can be used to explore and describe treatment techniques as well.

Sample Research Idea
A treatment manual for Qigong in the treatment of metastatic colon cancer based on a promising case report by Loh needs to be developed.

LEVEL SIX: TREATMENT EFFECTIVENESS STUDIES
Treatment effectiveness studies evaluate techniques of treatment, often by comparing the results of two different techniques in patients with similar patterns or diagnoses. The role of treatment effectiveness studies is to establish the benefit or lack thereof associated with the use of a particular treatment technique. Treatment effective studies are best done when the technique has been well described (i.e., after or concurrent with the publication of a technique study that specifies appropriate patients and clearly defines the characteristics or principles of the technique). The best-known and most respected type of treatment study in the biomedical community (but by no means the only type) is the RCT.

Toward the end of their lives, many OM scholar-practitioners published retrospective works comparing the effectiveness of various treatments. These include works by 20th century masters such as Qin Bowei and Jiao Shude, as well as well-known treatises such as Jiang Zhong-Jing’s Shang Han Lun and Sun
Si-Miao’s Qian Jin Yao Fang. These practitioners’ own continuity of experience serves as a kind of control; their lives and those of their patients were usually far more stable and standard than ours, increasing the consistency and internal validity of these reports to a level simply not achievable in modern industrial society. The obvious drawbacks of these studies are observer bias and lack of formal control groups.

While RCTs minimize observer bias and maximize internal validity, studying OM today through RCTs entails significant problems. Blinding the acupuncture practitioner is impossible, for instance, and patient blinding is possible only under limited circumstances and with acupuncture-naïve patients—a population that becomes, with every passing year, smaller and more self-selected for lack of interest in acupuncture. Furthermore, standardizing treatment between patients and between visits for the same patient appears necessary for study purposes but represents a marked departure from clinical practice. Good OM practitioners generally use palpation and measurement of pulse and other indicators (level two studies) to confirm that the treatment is working and that the diagnosis is correct (level four studies). With acupuncture, this assessment process often takes place in between needle placements, thus intermingling the measurement, diagnosis, and treatment processes. The vexing question of the placebo effect has been addressed with some success (level three studies). With acupuncture, this assessment process often takes place in between needle placements, thus intermingling the measurement, diagnosis, and treatment processes. The vexing question of the placebo effect has been addressed with some success (level three studies). With acupuncture, this assessment process often takes place in between needle placements, thus intermingling the measurement, diagnosis, and treatment processes. The vexing question of the placebo effect has been addressed with some success (level three studies).

Also necessary to RCT design are inclusion and exclusion criteria. Typically, patients are included in or excluded from studies on the basis of biomedical diagnoses. This results in a study group that appears homogeneous to biomedicine but is a random catchall of OM patterns, with a resultant need for flexi- bility in treatment that may invalidate the study from a biomedical point of view. What then is the principle by which we can sort out the inclusion and exclusion criteria for our RCT? Which patterns are similar enough from an OM point of view that the patient may be included and which so different that the patient must be excluded? (This problem bedevils any holistic approach to healthcare, for each human being taken as a whole defies categorization.) As of yet, we have no compelling answer to these questions because the theories of OM vis-à-vis Western research methodology and biomedical theory have not been deeply explored. For this problem, multiple level-one studies are needed.

Currently, when an RCT of OM fails to show results, it is open to interpretation as a failure of OM itself. However, if many successful level five studies are followed by failed RCTs, then the hypothesis must be considered that the RCT is not an effective tool for studying this aspect of OM. Whatever the reason for this ineffectiveness—even if it is that OM relies heavily on a placebo effect—a critical mass of successful level five studies constitutes reasonable grounds for moving to an experimental model that assesses effectiveness in a clinical setting.

Radically creative solutions may have value in exploring this state of affairs. One option could involve the use of qualitative studies to investigate practitioners’ and patients’ experiences and beliefs around which treatments are effective, which are not, how they are successful or unsuccessful, and why. Since public opinion is driving the acceptance of OM and CAM, why not capitalize on the involvement of the public through appropriate research designs? Other possible research designs include n-of-1 studies, in which a subject acts as his or her own control; treatment can be customized to the individual, patients can articulate preferences about treatment options, and the design can establish whether treatment is effective in that particular patient. Single-systems research designs also have a sample size of one and can be adapted to a wide variety of clinical problems. They can be used in patient care without significantly disrupting normal treatment.

The numerous resources for designing clinical trials include Wittmann and Walach, especially regarding designs that use natural groups, Stux and Birch (regarding standards of acupuncture treatment), and Bell et al for a whole-systems approach to trial design. More information about n-of-1 studies can be found in articles by Johnston and Mills, Backman and Harris, and Sung and Feldman, and about single-systems research design in Bloom. For an excellent review of options for placebo controls in randomized trials of acupuncture, see the work of Vickers and for advice regarding the selection of control treatments, Stux and Birch. Although considered to occupy the lower levels of evidence in evidence-based medicine, case reports and case series have great value in exploring the effectiveness of treatment in a holistic system such as OM.

Sample Research Ideas

(1) What treatments do experienced practitioners consider effective in the treatment of eczema, and how does their perception of effectiveness compare with Western published clinical studies? (2) Do patients with Lyme disease and candidiasis benefit when treated according to ancient notions of Gu syndrome (invasion of evil “worms”)?

LEVEL SEVEN: SYSTEMATIC REVIEWS

Systematic reviews draw together studies on the same topic to see if conclusions are thereby strengthened. When we search the published literature on OM to investigate a clinical question, we seldom find a completely satisfactory and reliably current answer. Where to find the information we seek, particularly since most of the literature is accessible only to those who read Chinese or can afford translation services? And once we find a publication addressing our issue of interest, how can we be sure that it is not contradicted or superseded by another publication? In such cases, our task can be greatly assisted by a review that surveys, summarizes, and critiques the published (and sometimes unpublished) material on a topic.

Review articles do exactly that. The best reviews, called systematic reviews, use explicit search strategies for finding studies and sources, use explicit inclusion and exclusion criteria to determine which studies and sources they will evaluate, give more weight to studies and sources that have better methodologies and are carried out better, and may combine statistical data from several studies to create stronger conclusions. In theory, the more systematic the review, the more unbiased its conclusions. In the form of meta-analyses, systematic reviews often evaluate...
completed studies of treatment for a particular condition “x.” Thus, the level of systematic reviews is placed here, following treatment effectiveness studies (level six). But systematic reviews can just as easily be used to evaluate and summarize foundation, measurement, group, pattern/diagnosis, and treatment technique studies.

Over the course of the past 2,000 years, OM has had ample opportunity to evaluate and synthesize its experience. Individual training has always emphasized studying a canon of foundational texts as well as the research and memoirs of past masters. In addition, at several junctures in Chinese medical history, regime change created the opportunity and motive for systematic review of the literature. For instance, the Yi Zong Jin Jian (Golden Mirror of Medical Orthodoxy) was completed under the newly installed Qing dynasty in 1742, and in the 1950s the early communist regime created Traditional Chinese Medicine.

Currently, we are at another of these junctures. As OM is received on foreign soil, how we organize our study of it will largely determine its future development. Communist thinking influenced the construction of Traditional Chinese Medicine, for instance, in the subordination of Five Element theory (based on the interaction of the elements water, fire, wood, metal, air, and earth, and rejected as feudal superstition) to the more obviously dialectical eight principle patterns (yin/yang, interior/exterior, repletion/vacuity, and cold/hot). One chief goal of the present paper is to provide a rational framework for minimizing similar distortions. If systematic reviews all ask the question, Is acupuncture safe and effective for “x” biomedically defined condition? we will indeed be greatly altering OM to suit our intellectual predispositions. We hope that level seven studies will be used consciously to advance the discipline of OM through self-study, as exemplified in the Standards for Reporting Controlled Trials of Acupuncture guidelines. Resources for conducting systematic reviews include Ezzo et al. (which also contains an example of a systematic review), Hart, and Oxman et al.

Sample Research Idea
With respect to a given herbal formula widely used over the centuries (eg, xiao chai hu tang), for what symptoms and OM patterns has this formula been used, and how has usage changed over time?

DISCUSSION
The levels presented in this paper could be used in a variety of ways to research clinical OM on its own terms: to determine if the requisite preparatory (lower level) studies have been completed before attempting to answer a specific (higher level) research question of interest; to identify areas of inquiry in OM that are particularly fertile or needed to make other research possible; and to stimulate the creative process of study design.

Figure 1 illustrates in general how the levels can build upon one another to strengthen our confidence in the findings related to a particular new treatment. This figure can also be used in research planning. If lower level studies have not been completed in answering a particular research question, they must be conducted before higher level studies can be fruitfully contemplated. The flowchart of Figure 2 sketches out one such scenario related to tongue color. Both Figures 1 and 2 show how the levels can be used in a circular, nonhierarchical manner.

Table 1 lists published OM articles that exemplify each level. Because the levels could be adapted for application to biomedicine as well as to systems of CAM such as Ayurveda and osteopathy, articles could be found for those systems as well. These systems proceed from different theoretical assumptions than OM, however, so that the levels for those systems could be
substantially different, especially with respect to the nature of diagnosis or assessment and its interaction with treatment.

Because OM is not only linear but also circular and reflexive, the levels can be blended and combined as needed and not just used in a hierarchical sequence, in the ways recommended by Walach et al. and Verhoef et al. For instance, patterns (level four) can also be diagnosed and validated through treatment (levels five and six); if the treatment specified for a particular pattern is efficacious, then the patient could be considered to manifest that pattern. Likewise, treatment effectiveness studies (level six) can shed much light on whether measurement instruments are working and how.

Many papers such as those cited in the introduction have created systematic approaches to research methodology with respect to CAM and even to acupuncture specifically. The levels presented here, however, were specifically designed to maintain the theoretical and clinical integrity of OM while providing a way to creatively apply Western research methodologies to furthering our knowledge of OM in itself and in relation to biomedicine. These levels were intended primarily for application to clinical questions and would thus hopefully be of use to clinicians and students as well as to experienced researchers. Broader views of research, which include laboratory studies, health services research, and economic analyses, are provided by the evidence house of Jonas, the balanced evidence hierarchy of Lewith et al., the five-data box model of Wittmann and Walach, and the system of proving, exploratory, and improving studies of Thomas and Fitter. We believe that the levels of this paper are compatible with these other approaches and able to be integrated with them, holding with Walach et al. that a multiplicity of methods used in complementary ways are needed, even to investigate the value of specific treatments. Indeed, within each level presented here, multiple methodologies are invoked as valuable in achieving the tasks of that level.

The question of how to fund such research in OM that—however interesting, novel, and needed—falls outside of the agendas for biomedically oriented funding sources is not a trivial one. The large-scale, double-blinded RCT (level six) is a research modality well suited to the testing of biomedical pharmaceutical agents, for instance, and the cost of such testing constitutes a reasonable investment on the part of the corporation holding the patent. In OM research, by contrast, funds are more likely to be granted by government or nonprofit entities with no expectation of a financial return—and are therefore relatively scarce, especially in relation to the high cost of adapting OM practice to a large-scale RCT scenario. The recent development of international CAM research networks may offer opportunities for lower-cost collaborative research.

Also, we hope that studies from levels one to five will be used to mitigate this logistical challenge in two ways. First, levels one to five can be invoked to generate lower-cost studies. The cost of one RCT could fund dozens of lower-level projects. Second, investigators could use lower-level studies to show solid OM foundation and demonstrate robustness of clinical protocols.

---

**Figure 2.** Selecting a research question related to tongue color using levels.
Rotchford and

Table 1. Published articles exemplifying the levels

<table>
<thead>
<tr>
<th>Level</th>
<th>Author</th>
<th>Goal of Article</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Park et al (^6^3)</td>
<td>Proposed a comprehensive definition for acupuncture in clinical research in light of the diversity of styles and the intertwined nature of treatment and diagnosis</td>
<td>Definitions and subject headings proposed could serve as a foundation for subsequent research communication</td>
</tr>
<tr>
<td>1</td>
<td>Zhang et al (^6^4)</td>
<td>Suggested that deqi (as measured by the verbal scale of Vincent et al, 1989) is a real phenomenon</td>
<td>Deqi was operationalized in a limited way (see level 2 below)</td>
</tr>
<tr>
<td>1</td>
<td>Vincent et al (^6^5)</td>
<td>Analyzed the existing literature on the relationship of treatment to diagnosis, OM diagnostic consistency, and the relationship between biomedical and OM diagnosis</td>
<td>Clarified diagnostic issues involved in relating these two medical systems</td>
</tr>
<tr>
<td>2</td>
<td>Yamashita et al (^6^6)</td>
<td>Developed a list of words to measure the sensations at the site of needling associated with acupuncture and the phenomenon of deqi</td>
<td>Deqi may also include phenomena related to muscular contractions, sensations from a larger area than the site of needling, and the perceptions of the acupuncturist as well as the patient</td>
</tr>
<tr>
<td>3</td>
<td>Hogeboom et al (^6^7)</td>
<td>Investigated the incidence of adverse reactions associated with acupuncture in a population of 391 consecutive outpatients</td>
<td>Lack of information about population demographics, diagnoses, and the practitioners’ treatment style limits generalizability</td>
</tr>
<tr>
<td>4</td>
<td>Chen et al (^6^8)</td>
<td>Developed a noninvasive placebo technique for acupuncture in the treatment of low back pain</td>
<td>Acupuncturists differed considerably in the specific patterns assessed in and treatments recommended for the same patient</td>
</tr>
<tr>
<td>5</td>
<td>Schiff et al (^6^9)</td>
<td>Explored the association between postmenopausal osteoporosis and the OM pattern of kidney-vacuity syndrome</td>
<td>Standardized criteria for kidney vacuity have not been translated into English</td>
</tr>
<tr>
<td>6</td>
<td>Porzio et al (^7^0)</td>
<td>Compared the OM patterns found by six acupuncturists in six patients with chronic low back pain</td>
<td>OM patterns were not reported in these patients</td>
</tr>
<tr>
<td>6</td>
<td>Beal and Nield-Anderson (^7^2)</td>
<td>Compared acupuncture, massage, and self-care education in treating chronic low back pain</td>
<td>This technique is for research, not clinical practice</td>
</tr>
<tr>
<td>7</td>
<td>Cherkin et al (^7^3)</td>
<td>Established the safety and efficacy of acupuncture in treating breast cancer patients’ menopause-related symptoms, which are frequently exacerbated in patients on tamoxifen</td>
<td>This single-group design could help in planning subsequent randomized controlled trials (RCT)</td>
</tr>
<tr>
<td>7</td>
<td>Van Tulder et al (^7^4)</td>
<td>Investigated studies of acupuncture in the management of acute and chronic low back pain, included trials with subjects who had nonspecific low back pain (LBP), and excluded those with LBP caused by specific pathologic entities</td>
<td>This was a pilot RCT of 11 patients in two groups.</td>
</tr>
<tr>
<td>7</td>
<td>McCulloch et al (^7^5)</td>
<td>Conducted a meta-analysis involving Chinese herbal medicine</td>
<td>Massage therapists were allowed to use trigger and pressure point techniques but acupuncturists were not allowed to use Chinese massage</td>
</tr>
</tbody>
</table>

The following limitations were identified for this model: Because it is oriented toward clinical research, it does not provide suggestions for the ways basic research and clinical research influence and inform each other. Because basic research is a complex, multifaceted topic in its own right, we deemed that introducing it here would add a baffling level of complexity for the great majority of students, clinicians, and clinical faculty whom we hope to assist. Furthermore, within the scope of this article we...
can only offer the barest suggestions about how to design and carry out research in each level. Nonetheless, we hope that within its limited sphere, this model of levels will serve an invigorating and useful function in clinical research in OM.

CONCLUSION
In this foundation study (level one), we have presented a way of approaching research that hopefully makes sense to OM practitioners, students, and teachers in light of the way they approach patients and the theories they use. We also hope that the model presented here will help generate viable research ideas for investigating both OM theory and practice and the issues surrounding the integration of OM and biomedicine. We believe that this approach could be fruitfully applied to other treatment modalities in the fields of alternative, complementary, and holistic healthcare, such as osteopathic manipulation and Ayurveda. Finally, we hope that by providing a culturally sensitive framework for the exploration of medical care, the levels will facilitate biomedicine in rationally reevaluating its own conceptual foundations.

REFERENCES