Withdrawal/Redaction Sheet
Clinton Library

DOCUMENT NO. AND TYPE	SUBJECT/TITLE	DATE	RESTRICTION
001. resume	Robert Blanks (partial) (1 page)	nd	b(6)

COLLECTION:
Federal Records
White House Commission on Complementary and Alternative Medicine Policy [WHCCAMP]

OA/Box Number: 43232

FOLDER TITLE:
WHCCAM Commissioner Binder, Day 3, May 16, 2001 [2]

2011-0568-S
kc813

RESTRICTION CODES

Presidential Records Act - [44 U.S.C. 2204(a)]

P1 National Security Classified Information [(a)(1) of the PRA]
P2 Relating to the appointment to Federal office [(a)(2) of the PRA]
P3 Release would violate a Federal statute [(a)(3) of the PRA]
P4 Release would disclose trade secrets or confidential commercial or financial information [(a)(4) of the PRA]
P5 Release would disclose confidential advice between the President and his advisors, or between such advisors [a)(5) of the PRA]
P6 Release would constitute a clearly unwarranted invasion of personal privacy [(a)(6) of the PRA]

C. Closed in accordance with restrictions contained in donor's deed of gift.
PRM. Personal record misfile defined in accordance with 44 U.S.C. 2201(3).
RR. Document will be reviewed upon request.

Freedom of Information Act - [5 U.S.C. 552(b)]

b(1) National security classified information [(b)(1) of the FOIA]
b(2) Release would disclose internal personnel rules and practices of an agency [(b)(2) of the FOIA]
b(3) Release would violate a Federal statute [(b)(3) of the FOIA]
b(4) Release would disclose trade secrets or confidential or financial information [(b)(4) of the FOIA]
b(6) Release would constitute a clearly unwarranted invasion of personal privacy [(b)(6) of the FOIA]
b(7) Release would disclose information compiled for law enforcement purposes [(b)(7) of the FOIA]
b(8) Release would disclose information concerning the regulation of financial institutions [(b)(8) of the FOIA]
b(9) Release would disclose geological or geophysical information concerning wells [(b)(9) of the FOIA]

Complementary Therapies in Medicine (1996) 4, 198–201

FOR DEBATE

Old myths given new voice. The Nuffield Report: *Researching and Evaluating Complementary Therapies: The State of the Debate*

A. Vickers

Research Council for Complementary Medicine, London

SUMMARY. In May 1995, the Nuffield Institute for Health published a report on the methodology of complementary medicine research. Entitled *Researching and Evaluating Complementary Therapies: The State of the Debate*, the report revives a number of misconceptions about conventional research techniques and their application to complementary therapies.

MYTH 1: RESEARCH AND RESEARCH ON EFFECTIVENESS ARE THE SAME THING

Throughout the report, the authors used the word 'research' as a synonym for 'research on effectiveness'. This is explicitly stated on page 24 '... no attempt has been made to extend the discussion beyond a concern with clinical effectiveness'. Yet just a few lines down, under a section entitled *Researching Complementary Therapies*, the authors discuss '... issues which need to be considered when undertaking *research* into complementary therapies,' [my emphasis]. This careless use of terminology makes it difficult to evaluate claims such as that the randomized controlled trial (RCT) is the 'gold standard' for medical research. If medical research is inaccurately seen as synonymous with effectiveness research, this may be true. However, if research is seen more generally as systematic inquiry, this assertion is simply false. Researchers use a wide variety of different techniques to answer a range of different questions. It is open to doubt whether any worker would try to use an RCT to answer questions such as, what sorts of touch do nurses use? what is patients' experience of pain and its treatment? or what is the prognosis for premature infants?

MYTH 2: CHOICE OF RESEARCH METHODOLOGY DEPENDS ON THE THERAPY RATHER THAN THE QUESTION UNDER INVESTIGATION

Perhaps the key claim of the report is that 'The choice of research design should be made on the basis of which methodology is most appropriate to researching the therapy under investigation' (p. 7). It is difficult to understand how this strategy could be put into practice. Researchers normally choose research designs depending on the question they intend to ask. The advantage of this approach can be seen from the following example.

Imagine that a researcher wished to ask three questions about a certain therapy: (i) how many people used the therapy last year? (ii) in what different ways do practitioners use touch during a consultation? and (iii) does the addition of the therapy to a standard package of care improve outcome in chronic low back pain? Using the question-oriented approach, the choice of research design is relatively clear: the researcher would use some form of quantitative population-based survey, a qualitative research approach, and a clinical trial respectively. If using the strategy suggested by the Nuffield report, the researcher would presumably have to determine the most appropriate methodology for the therapy and then try to use that methodology to answer all three questions. This might entail a researcher having to use an RCT to find out how physiotherapists use touch (physiotherapy is a form of conventional medicine where the RCT is said to rule supreme). Likewise, some form of qualitative research or case-study approach (said to be characteristic of complementary medicine) would need to be used to find out how many people visit practitioners of osteopathy each year.

A particularly good example of the author's confusion over the links between questions and research designs is the authors' description of Taylor-Reilly et al's study of homoeopathy.[1] This is described as overlooking such considerations as the authenticity of diagnosis and appropriateness of standardizing treatment. A focus on the research question helps clarify the issue. These researchers did not set out to examine

Andrew Vickers, Information and Resources Director, Research Council for Complementary Medicine, 60 Great Ormond Street, London WC1N 3JF, UK.

the effectiveness of homoeopathy as practised. They were instead attempting to answer the question 'Can a homoeopathic remedy show effects above and beyond a placebo response?'. Many questions need answering in health research. Clarity as to the question in hand is vital if we are to design, evaluate and use research effectively.

MYTH 3: TECHNIQUES FOUND IN SOME RANDOMIZED CONTROLLED TRIALS ARE AN INHERENT PART OF RANDOMIZED CONTROLLED TRIAL METHODOLOGY

On page 5, under a section entitled *The RCT Protocol*, the authors list a number of methodological techniques. Though all of these techniques are found in at least some RCTs, many are not an inherent part of RCT methodology. In other words, it is perfectly possible to dispense with these techniques and conduct a methodologically sound RCT. It is claimed that RCTs involve the following features:

- *The identification of a hypothesis which is 'located within agreed theoretical knowledge in biomedicine'*. The mechanism of action of many therapies is poorly understood, some more so than others. This has not, however, prevented their examination by RCT methodology. RCTs have been conducted on therapies as diverse as prayer,[2] distant healing,[3] homoeopathy,[1] cell therapy[4] and reflexology;[5] and the results published in mainstream journals. None of these therapies can be said to involve agreed biomedical knowledge.
- *A focus on a 'defined diagnostic group/disease'*. As long as researchers set reasonable eligibility and exclusion criteria, there is no necessity to focus on a particular disease. For example, Frederikson's[6] study of a patient education leaflet included a spectrum of patients and measured doctor ratings of communication quality as an outcome.
- *The use of a standardised treatment*. This claim is made most explicitly on page 11 when the authors state that 'the scientific method' leads to 'the transformation of a flexible and individualized therapy into a packaged formula'. As it happens, there is absolutely no requirement that the treatment in RCTs should be standardized. This would, for example, have rendered impossible most trials of surgery, where surgeons need to make treatment decisions depending on the particular characteristics of each patient. Likewise, many RCTs of complementary therapies have been undertaken where practitioners have been free to practise 'as normal' and where only the most minor constraints were placed on the specific treatment techniques used and on the timing frequency and duration of treatment. See, for example, the trials of Meade[7] and de Lange de Klerk.[8]
- *The use of the double-blind method and placebo controls*. For many, perhaps even most medical techniques, it is not feasible to use either placebos or double-blinding. Examples include surgery, physiotherapy, speech therapy and screening procedures, as well as many complementary therapies. Again, RCTs have been used to investigate such techniques.
- *The use of 'medically defined (objective) outcome measures avoids any reliance on subjective patient reports of their condition' (p. 6)*. This problem has, as the authors note, been recognized in conventional research. As a result, possibly the majority of clinical trials now involve patient reports of subjective experiences such as pain, mood, nausea, or disability.

MYTH 4: CONVENTIONAL AND COMPLEMENTARY MEDICINE ARE DISCRETE HOMOGENOUS ENTITIES

One example of this misconception, which runs throughout the report, is the author's claim that complementary therapies reject 'biomedicine's focus on disease in terms of biochemical processes and their breakdown, the emphasis on specific aetiology and the mind – body dualism. This is replaced by a systems approach to health and healing' (p. 8). There are a number of problems with this claim. First, there is enormous diversity within both conventional and complementary medicine. How many generalizations hold for both plastic surgery and nursing? How many for chiropractic and yoga? Furthermore, as the authors themselves caution, there are often radical differences between practitioners within a therapy. Also problematic is the implicit assumption that complementary and conventional medicine are discrete (i.e., non-overlapping). In the above example, we are encouraged to think that there is no element of mind–body dualism in one form of medicine and nothing of a systems approach in the other.

Most fundamentally, no evidence is presented to support the alleged characteristics of complementary and conventional medicine other than a small number of complementary practitioners' say-so. The authors could, for example, provide sociological data on conventional clinicians' attitudes and behaviours and compare this with those of their complementary counterparts. Rather than doing so, they develop caricatures of both, and thus do scant justice to the range and diversity of medical practice. In what way does hospice care involve an emphasis on specific aetiology? Is it impossible for a good GP to be holistic? In what ways do those chiropractors who believe subluxations are responsible for all illness represent

an escape from specific aetiology and mind–body dualism? Such questions remain unanswered. Instead, the authors present, for example, a quote from a *British Medical Journal* editorial from 1980 complaining about the lack of scientific rationale for chiropractic. It is unlikely that such an editorial is a sociologically accurate reflection of medical practice. Similarly, it is unlikely that the methodological literature is an accurate reflection of what complementary practitioners actually believe.

MYTH 5: COMPLEMENTARY PRACTITIONERS HAVE UNDERTAKEN QUALITATIVE RESEARCH AND THIS IS OF VALUE FOR DETERMINING WHETHER A THERAPY IS EFFECTIVE

It is claimed that 'more quantitative methods' are 'preferred by orthodox medicine' and 'more qualitative methods' are 'associated with complementary therapies' (p. 24). No evidence is presented to support this claim. Not one example of qualitative research in complementary medicine is quoted in the report. Though the case studies which are found in many complementary journals might be considered qualitative in nature, it is doubtful that any meet even a small number of the methodological criteria used to assess the rigour of such research.[9]

The authors also complain about the low status of qualitative research in the hierarchy of evidence (p. 24) and recommend that 'research... give greater credence to the use of qualitative methods' (p. 24). It is not generally thought that qualitative research can, by itself, assess questions of effectiveness. Among other roles, qualitative methods can play an important part in determining the questions asked in quantitative trials and in helping to implement their results. However, they are not usually thought to produce direct evidence of effectiveness. The authors of the Nuffield report do not explain how qualitative methods could play such a role in complementary medicine.

MYTH 6: IF YOU WANT TO LEARN ABOUT METHODOLOGY, TURN TO THE METHODOLOGICAL LITERATURE

It is interesting to speculate on why the authors of the Nuffield report have made the aforementioned errors. The main reason seems to be that they have consulted the methodological literature rather than the primary research, i.e. actual studies on patients or practitioners. Of the 75 or so references quoted in the report, only 7 or 8 refer to primary research papers, none of which was published in the past five years.

A telling part of the report starts with the statement that 'a central thread running through the literature on complementary therapies, compared with studies of orthodox medicine... [is] the importance of integrating "patient acceptability" as a criterion of outcomes.' It concludes with a list of questions 'which take subjective experience... seriously' including 'What recognition is given to patient preferences in treatment?' and 'What is the level of compliance by patients?' Though this may be an accurate description of the opinions found in the methodological literature, it is not a good reflection of things as they really are. In conventional medicine there have been numerous studies of patient preference (e.g. Fallowfield's[10] work on choice in breast cancer) and compliance (see, for example, the extensive overview by Haynes et al[11]). The same does not appear to be true in complementary medicine; in fact, it is debatable whether any compliance or preference studies have been conducted whatsoever.

In relying so heavily on the methodological literature, the authors of the Nuffield report failed to use another useful resource: active researchers. Despite the claim that the report represents 'the state of the debate' no attempts seem to have been made to collaborate with individuals at research centres such as the Centre for the Study of Complementary Medicine in Southampton, UK, the Research Council for Complementary Medicine, the Royal London or Glasgow Homoeopathic Hospitals, UK, the Office of Alternative Medicine in the USA or the Centre for Complementary Health Studies at the University of Exeter, UK. The authors do not describe the methods they used to assess 'the state of the debate' and so it is difficult to know whether they have done so accurately.

In conclusion, by relying on the methodological literature of the mid '80s and early '90s, the Nuffield report revives a number of misconceptions about complementary medicine research. Those interested in conducting research would do well to avoid making this same mistake and to turn instead to the primary literature and the advice of those active in the field.

REFERENCES

1. Taylor Reilly D, McSharry C, Taylor MA, Aitchison T. Is homoeopathy a placebo response? Controlled trial of homoeopathic potency, with pollen in hayfever as a model. Lancet 1986; 8512 (18 Oct): 881–885.
2. Byrd R. Positive therapeutic effects of intercessory prayer in a coronary care unit population. South Med J 1988; 91(7): 826–829.
3. Beutler JJ, Attevelt JT, Schouten SA et al. Paranormal healing and hypertension. BMJ 1988; 296 (28 May): 1491–1494.
4. Black DB, Kato JG, Walker GW. A study of improvement in mentally retarded children accruing from Siccacell therapy. Am J Mental Defic 1966; 70(4): 499–508.
5. Oleson T, Flocco W. Randomized controlled study of premenstrual symptoms treated with ear, hand, and foot reflexology. Obst Gynec 1993; 82(6): 906–911.
6. Frederikson LG, Bull PE. Evaluation of a patient education leaflet designed to improve communication in medical consultations. Patient Educ Couns 1995; 25(1): 51–57.
7. Meade TW, Dyer S, Browne W et al. Low back pain of mechanical origin: randomised comparison of chiropractic and hospital outpatient treatment. BMJ 1990; 300 (2 June): 1431–1437.

8. de Lange de Klerk ES, Blommers J, Kuik DJ et al. Effect of homoeopathic medicines on daily burden of symptoms in children with recurrent upper respiratory tract infections. BMJ 1994; 309(6965):1329–1332.
9. Mays N, Pope C. Rigour and qualitative research. BMJ 1995; 311: 109–112.
10. Fallowfield LJ, Hall A, Maguire GP, Baum M. Psychological outcomes of different treatment policies in women with early breast cancer outside a clinical trial. BMJ 1990; 301(6752):575–580.
11. Haynes RB, Taylor DW, Sackett DL. Compliance in health care. Baltimore, MD: John Hopkins University Press, 1979.

Copies of the report may be obtained from the Nuffield Institute for Health, 71–75 Clarendon Road, Leeds LS2 9PL (Tel: 0113 2336983)

Vol. 13 Winter 1997

International Journal of

TECHNOLOGY ASSESSMENT IN HEALTH CARE

CAMBRIDGE
UNIVERSITY PRESS

International Journal of Technology Assessment in Health Care, 13:1 (1997), 111–121.

HOW SHOULD WE RESEARCH UNCONVENTIONAL THERAPIES?

A Panel Report from the Conference on Complementary and Alternative Medicine Research Methodology, National Institutes of Health

Andrew Vickers
Research Council for Complementary Medicine

Barrie Cassileth
University of North Carolina

Edzard Ernst
University of Exeter

Peter Fisher
Royal London Homeopathic Hospital

Peter Goldman
Harvard School of Public Health

Wayne Jonas
National Institutes of Health

Sung-Keel Kang
College of Oriental Medicine

George Lewith
University of Southhampton

Ken Schulz
Centers for Disease Control and Prevention

Chris Silagy
Flinders University of South Australia

Abstract

Research in unconventional medicine requires a number of different questions to build up a "mosaic" of evidence. Choice of research design depends on the question being asked and is independent of the therapy under investigation. Despite the doubts of some practitioners, randomized trials are of value for determining certain questions in alternative medicine.

The Office of Alternative Medicine (OAM) was established in 1991 as part of the federally funded U. S. National Institutes of Health. One aspect of the office's charge is to promote and fund research on unconventional forms of health care. Practitioners of these techniques often refer to understandings of the world that differ significantly from those usually found in biomedicine. As a result, they have raised fundamental questions about the validity of standard research methods. Is it appropriate to use conventional research techniques on unconventional therapies? Can a research strategy developed in a biomedical context be used to investigate a medical system based on a different "world view"?

In response to these challenges, the OAM sponsored a series of conferences to examine the appropriateness of conventional research methodologies for unconventional medicine. These conferences analyzed key research principles, and their findings may be of interest to a wider audience. The following report of the panel outlines how different types of evidence can contribute to an understanding of a treatment.

The panel was charged with determining how different types of evidence contribute to making an assessment of therapies described as complementary or alternative medicine (CAM). We decided early that "types of evidence" was related to "types of questions." Research is a tool for answering questions: the question therefore comes first. Moreover, classic problems in the design and interpretation of CAM research have often been solved by close attention to the question being asked in research. In the past, practitioners have asked, "What research designs are appropriate for CAM?" (23). A better strategy seems to be to identify the particular question being asked and to fit the research design to the question. Similarly, debates over the meaning of certain studies often turn on the question being asked in the study.

The panel decided to try to answer four main questions:

1. What is a good question?
2. How can questions be matched with research designs?
3. What is a strategic approach to research?
4. How can inappropriate interpretations of trials be avoided?

However, we thought that it would be worth spelling out our assumptions first:

- We have an incomplete understanding of the world: knowledge evolves and develops.
- Research methods may evolve and change as we continue to learn about the world.
- Research is always feasible—and essential—regardless of the therapy under consideration.
- Research rarely provides unequivocal answers: there is often room for reasonable people to disagree, within limits, about the interpretation of evidence. Nonetheless, some statements made about health are incorrect, and some practices are not in the best interests of patients.
- Good research aims to minimize the effects of bias, chance variation, and confounding.
- Our priority is research that investigates whether treatments do more good than harm.

On a semantic note, it was agreed that "questions" subsumes the term hypotheses. It was also agreed that "questions" would be defined as those associated with individual studies. Some questions need to be broken down into stages. For example, the question "What is the role of acupuncture in the treatment of pain?" or "What are the effects of patient expectations on outcome?" will not be resolved by a single research study. Researchers need to break down large, global questions into manageable stages

of a series of questions, each associated with a single study. Research should be conducted one study at a time.

WHAT IS A GOOD QUESTION?

Designing a research question requires knowledge and understanding of the subject under study. Just as we would not expect a CAM practitioner to be able to design a study of surgery without consulting a surgeon, we should not expect a researcher without experience with CAM to design a suitable trial. There are cases in which an insufficient understanding of CAM and/or research methodology has led to poor quality research. For example, in one early trial of homoeopathy (28) it has been claimed (12) that the criteria used to select remedies were inappropriately applied. Similarly, it has also been suggested that the negative result of some studies of acupuncture (17) stemmed from the incorrect assumption that needling incorrect acupuncture points is equivalent to placebo. We would recommend that CAM research should be a team effort, involving experts both in therapy and in research methodology.

There are two general criteria for evaluating a question: (*1*) Is it answerable? and (*2*) Is it important?

Is it answerable?

A question is likely to be answerable if it is explicit, focused, and practicable.

Explicitness. It is well known that the same word can mean different things to different people. For example, the meaning of the word "effective" in the question, "Is acupuncture an effective treatment for asthma?" is probably different for patients, who would be concerned whether treatment would significantly improve their quality of life, than for a scientist, who might, for instance, be concerned with whether acupuncture had an effect on bronchial reactivity beyond that of placebo.

Focus. Research almost always turns out to be more complicated than expected. Lack of focus is a common problem in all types of research, with researchers often attempting to answer a number of questions at a time. For example, in one proposed study on breast cancer, the authors planned both to document the prevalence of particular beliefs among patients and to examine the influence of the acceptance of CAM on the range and variety of discourses between patients and caregivers. It was also proposed that effects of particular types of discussion should be examined with a view to providing suggestions for improving verbal interactions between patients and caregivers. Such complex designs, involving multiple endpoints and outcome measures, should be avoided by all but the most experienced researchers.

Practicability. A particular question can be impractical for a number of reasons, including resources or patient numbers. Suitable research tools also need to be available. For example, the question of whether fennel oil strengthens the "heart chakra" (5) does not appear to be answerable at the time of writing because, as far we understand, there is no reliable means by which such an effect could be measured (which is not to say that a suitable method could not be devised).

Is It Important?

A question is important if the answer would lead, directly or indirectly, to the alleviation or prevention of suffering; it is of personal importance to the researcher and it is of importance to a social group. Alleviation or prevention of suffering is a fairly

universal criterion for all health practitioners and researchers. Personal importance is a consideration on the grounds that research requires a significant expenditure of time and effort and is unlikely to be brought to a successful conclusion unless the researcher has a commitment to that end. In addition, social groups, from countries to practitioner groups to individual hospitals, can and do set priorities for research. The criteria traditionally used during priority-setting exercises in conventional medicine are: the burden of illness; the cost to society; the knowledge gap; and the adequacy of current therapy. In the case of CAM, the extent to which the therapy is used by the public should also be taken into consideration.

HOW CAN QUESTIONS BE MATCHED WITH RESEARCH DESIGNS?

Matching questions with research designs is a two-stage process. First, it is necessary to define the general type of question and from there deduce an appropriate methodology. As a simple example, a question about public utilization of CAM would require observational, sociological research, whereas a question about the attribution of an effect to a clinical intervention would require a clinical trial.

Second, it is necessary to match particular aspects of questions with particular aspects of studies. For example, whether sociological research on public utilization employs a quantitative or qualitative methodology would depend on whether the question could be answered only in words (e.g., "What prompts patients' first approach to a CAM practitioner?") or whether numbers would be an important part of the answer (e.g., "What percentage of new CAM patients claim to be unsatisfied with conventional approaches?"). Similarly, whether the control group in a clinical trial received standard care, no intervention, or a placebo would depend on whether the question concerned a comparison (e.g., "Is new drug *x* superior to standard antiemetics?"), the existence of an effect (e.g., "Is acupuncture of benefit in asthma?"), or the nature of that effect (e.g., "Does acupuncture have specific effects in the treatment of asthma?"). A number of general categories of question, and examples of each type, are given in Table 1.

Attributing Cause and Clinical Effect

We felt that matching a general category of question with a research design was relatively straightforward and uncontentious, with the exception of attributing cause and clinical effect. This is possibly the topic of primary interest in CAM research: the question, "Does it work?" means in essence, does a CAM therapy cause a clinically worthwhile change in a patient?

Conventionally, the randomized controlled trial (RCT) is considered to be the "gold standard" for attributing cause and effect. The basic argument for the RCT is that it is important to judge the worth of a therapy in comparison with something else (health practitioners usually have a number of choices, including doing nothing) and that only by randomizing can we ensure that we are comparing like groups. There is a considerable literature supporting the use of RCTs. This literature suggests that RCTs are able to distinguish ineffective from effective therapy more reliably than other types of study (22;24).

Some practitioners of CAM, among others, have criticized the RCT as a means of investigating effectiveness or claimed that, in general, RCTs are infeasible or an inappropriate means of answering questions about the clinical effectiveness of a CAM therapy. Arguments about the RCT often seem to conflate the separate issues

Table 1. Categories of Questions in Complementary Medicine Research

Category of question	Examples
Attributing cause and clinical effect	"Does homeopathically prepared grass pollen reduce symptoms of hay fever more than a placebo?" "Is polypharmacy more effective than a single remedy approach in the homeopathic treatment of chronic hay fever?"
What happens in clinical practice?	"What is the cost effectiveness of adding homeopathic treatment to a standard care package in hay fever?" "What are the patterns of cross-referral between conventional and CAM practitioners in a multidisciplinary pain clinic?" "How common are serious neurological complications following chiropractic cervical manipulation?"
What do people do?	"How many people visit practitioners of CAM each year?" "What do patients tell their primary care physician about usage of CAM?" "How many nurses practice complementary medicine?"
What do people believe and how do they explain it?	"What do nurses believe about therapeutic touch?" "What is the patient's experience of the acupuncture consultation?"
By what mechanisms does a therapy work?	"What are the effects of needling the Hoku point on the production of endogenous opiates?" "Does peppermint oil reduce histamine-induced contractions of tracheal smooth muscle?"
Does something proposed in a therapy actually exist?	"Does homoeopathically prepared copper ameliorate the effects of copper poisoning in a plant model?" "Can acupuncture points be distinguished from non-acupuncture points by measuring the electrical resistance of the skin?"
Is a diagnostic or prognostic test accurate?	"How sensitive and specific is detection of gall bladder disease by examining photos of the iris?" "Is tongue diagnosis reliable?"

of scientific rigor and practical feasibility. For example, it is pointed out that RCTs would be the wrong sort of research for examining whether early sexual activity causes cervical cancer. On examination, this is for practical reasons: it would be unethical and infeasible to randomize teenage girls to different levels of sexual activity.

We thought it would be worth conducting a thought experiment to tease out the various problems thought to be associated with the RCT. In the thought experiment, we imagined a place where there were no practical constraints on research. We would have as much time and money and as many patients as we needed. Subjects would always do exactly what they were told and there were no ethical limitations on what they could be asked to do. We would also be freed from the obligation to provide care. The question posed was: if there were no constraints on time, patient numbers, patient compliance, ethics, or medical practice, and we wished to investigate cause and clinical effect, would we always conduct an RCT? Noting certain assumptions, we concluded, "yes." In other words, the only reasons not to conduct an RCT to attribute cause and clinical effect are practical: there are no theoretical arguments against the RCT for this purpose.

Assumptions Underpinning the RCT

We are aware that our view of the RCT as the best existing means of attributing cause and effect is predicated on three assumptions: (*1*) cause precedes effect; (*2*) mental beliefs do not influence chance events; and (*3*) the beliefs of the researcher do not have unmediated physical effects.

The third assumption is worth discussing in more detail. First, it is worth emphasizing the importance of the word "unmediated." Observer effects are well known in many areas of research: in sociology, for example, the notion that people change their behavior when they are being watched (the so-called Hawthorne effect) is a key methodological insight. However, this effect is mediated (for example, by expectation and social pressure) and can be subject to controls. Instead, we are dealing here with the suggestion (made, for example, by parapsychologists and some healers) that the mere existence of a researcher, with beliefs one way or another, can alter the outcome of an experiment. We are assuming that this does not occur.

We would also like to stress that our assumption that beliefs do not have unmediated effects applies only to the researcher. The assumption does not rule out research on modalities, such as prayer, in which the belief or will of one individual is said to cause a change in a patient that is unmediated (at least by nonsupernatural forces). The beliefs of a healer might have unmediated effects on a patient; we will assume, however, that the beliefs of the researcher do not.

These three assumptions seem reasonable: they are the sort of assumptions we use go about our daily lives. Nonetheless, these assumptions are open to empirical disconfirmation. There are those who would claim that this has already occurred, perhaps quoting parapsychological or quantum mechanical research. However, these effects are likely to be small, particularly in comparison to those resulting from factors such as bias, which is known to have a large effect on the outcome of poorly conducted research (24).

In summary, matching a question to a research design requires carefully deciding the general type of question to be studied, and from there deducing an appropriate methodology. Possibly the most important, and contentious, type of question in CAM research is the attribution of cause and effect. Given certain reasonable assumptions, the RCT is the most reliable research design for answering this type of question,

even in cases where there are complex webs of causation. However, there are practical and ethical considerations that may make it infeasible to conduct an RCT.

Numerous forms of study can be used to attribute cause and effect if an RCT cannot be conducted. These include case-control studies, cohort studies, and nonrandomized comparisons using, for example, historical controls, case series, etc. Although useful, these studies are not as rigorous as the RCT and the data that are produced are not as reliable. There are numerous examples where an intervention was found to be effective by a case series and then shown to be ineffective by the RCT. Fuller (13) reported on a case series of 194 smokers treated with acupuncture. Forty-one percent cessation rates were reported at 6 months. In a subsequent randomized trial (3), however, cessation rates were not only much lower (22% at 1 month) but no difference was found between real and placebo acupuncture. Dale and Cornwell (4) reported the results of both an uncontrolled trial, in which 85% of women given lavender oil following labor stated that it was of benefit, and a subsequent RCT in which no difference was found between true lavender oil, a synthetic lavender substitute, and an inactive fragrance compound. Similar examples can be found in conventional medicine. Cardiologists once believed that it was worth giving certain antiarrhythmic agents to patients with mild or asymptomatic arrhythmias. An RCT known as the Cardiac Arrhythmia Suppression Trial (CAST) (11), however, demonstrated that use of such agents led to higher mortality than placebo. In the case of attributing cause and clinical effect, the RCT seems to be the best way of "preventing [us] deceiving [ourselves] in regard to [our] creatively formed subjective hunches" (21).

This brings us to a very important point: for one particular category of research question (attributing cause and effect), the RCT is the best form of study yet developed and therefore part of an inherent hierarchy of types of evidence. However, there is no hierarchy among categories of question. If the question is about what happens in clinical practice or about what patients are doing and why, the RCT is unlikely to be the design of choice. A variety of different types of questions (Table I) must be asked and different types of evidence used to create a mosaic picture for making judgments about clinical practice. Each different type of evidence has its particular role, strengths, relevance, and limitations. An efficient research strategy would provide a balanced approach, filling in the knowledge about all types of general questions.

It is worth restating that the RCT is not the only valid type of research. A number of different types of research methods are valid. Each is appropriate for answering particular types of questions. Just as qualitative research is the method of choice for investigating what people believe about a novel subject, so the RCT is the method of choice for attributing cause and effect, particularly with regard to the effects of a therapy.

WHAT IS A STRATEGIC APPROACH TO RESEARCH?

It is only appropriate to ask certain questions once the answers to other questions are known. The order in which research questions are approached can sometimes be illogical. One common mistake has been to jump immediately to placebo-controlled RCTs of a standard therapeutic regimen without first having good preliminary evidence (e.g., from a pilot study or case series) that the regimen is of benefit. The study of homoeopathy in osteoarthritis by Shipley et al. (28) is a case in point: the practitioners involved were limited to the use of a single remedy. No pilot study or case series had demonstrated that this remedy could be used to treat osteoarthritis effectively.

Common failures to develop a strategic approach to research include:

- Using untested interventions in definitive trials;
- Researching the mechanisms of a nonpharmacological therapy before establishing whether it is of clinical benefit;
- Developing outcome measures at the same time as evaluating a therapy;
- Replication of negative research; and
- Failure to balance theoretical and practical questions.

Some model examples of good strategic approaches to research are useful. Dundee, an anesthetist with considerable experience with testing antiemetic agents, observed the use of the P6 acupuncture point to treat morning sickness in China and wondered whether this technique might help to control postoperative vomiting. He first conducted a nonblinded, nonrandomized, pilot study comparing P6 acupuncture plus standard antiemetic medication to antiemetics alone (6). The results were positive for P6, thus validating the study model and paving the way for more rigorous research. Dundee then conducted a number of blinded, placebo-controlled trials (6;7;8) before undertaking open trials aimed at determining the most effective and acceptable methods of P6 stimulation (9;10;19). Koes and colleagues (15) undertook a comprehensive systematic review of manipulation for low back pain. Significant flaws in published research were identified before a new clinical study was designed (16). Atherton and coworkers (1) conducted an open assessment of a Chinese herbal preparation in eczema, which, when found to be beneficial, was more rigorously tested in two randomized controlled trials (25;27). To gain additional information, the patients taking the therapy were then followed for an additional 2 years to assess long-term safety and efficacy (26).

HOW CAN INAPPROPRIATE INTERPRETATIONS OF TRIALS BE AVOIDED?

Interpretation of the results of research is fundamental to the scientific process. Inappropriate interpretation can render studies worthless and has been a problem for CAM.

Evidence that an agent has pharmacological effects in animal or in vitro models does not imply that it is of clinical value. Aromatherapists have sometimes made this mistake in trying to evaluate the clinical effects of herbal products on the basis of laboratory evidence (20).

Statistical significance and clinical significance differ: statistical differences in favor of a treatment do not imply that it is necessarily of clinical value. On the other hand, clinical significance is not always relevant. One criticism of Reilly and colleagues' hay fever trial (29) is that the use of homoeopathy led to only small improvements in symptoms. However, they only aimed to investigate one particular question "Is homeopathy a placebo?," and the clinical value of homoeopathy was not at issue.

It is often inadvisable to make general statements about efficacy on the basis of trials of particular techniques with particular groups of patients. One negative trial of two homoeopathic remedies for the resumption of ileus after surgery (18) provoked a reaction against the therapy as a whole in the French press. Similarly, practitioners sometimes use trials to vindicate and support their practices, even when the treatment used in the trial was unlike the one they use in everyday practice.

Research that contains methodological flaws cannot be used as a basis for knowledge claims. Papers should be critically appraised and interpreted in the light of any methodological shortcomings.

Literature reviews have sometimes confused "no evidence of effect" with "evidence of no effect," a particular problem when investigating adverse effects of treatment. Moreover, if a study fails to demonstrate an effect, this does not mean that an effect does not exist, only that it could not be found by the study as carried out. One common reason for failing to find an effect is that the sample size is so small that the trial lacks sufficient power to show a statistically significant difference between the treatment and control groups.

The research evidence on a particular topic needs to be evaluated as a whole. Authors often cite only research that supports their own particular viewpoint. An example is the claim that mental attitudes increase survival time in cancer. Texts aimed at the general public often contain reference to a variety of positive studies but fail to mention trials that fail to show an effect of psychological factors or psychosocial interventions (2;14).

Single studies are rarely dependable; research should generally be replicated. Moreover, subgroup analyses only generate hypotheses. It has not been uncommon for practitioners to use analyses of subsets of data to support their practices.

Reviewing these errors, we felt that avoiding inappropriate interpretations was a matter of following four steps:

1. Interpreting the study in terms of the question: was the design appropriate?
2. Evaluating the study in terms of other relevant evidence: what other data can be brought to bear?
3. Assessing the methodological rigor by which evidence was obtained: was the research properly executed?
4. Judging the generalizability of the study results: to what situations can we apply the results?

More generally, we felt that appropriate interpretation is an inherent part of science and therefore depends on fostering a critical attitude toward the evaluation of evidence. The pitfalls in interpreting evidence should be an integral part of courses in research methodology. To practice evidence-based health care, we must understand the means by which health care can be based upon evidence.

QUESTION-ORIENTED RESEARCH: DOES IT WORK?

One of the interesting aspects of the conference was the opportunity to evaluate our approach to research through discussion with CAM practitioners. We were confronted by a number of practitioners who wished to conduct research, but who felt overwhelmed by its apparent complexity — and suspicious of techniques thought only to be appropriate to conventional medicine. We encouraged them to concentrate on a single question at a time and to match the question to a research design. Practitioners found that they felt comfortable with this process and realized that research would not necessarily be "unfair" to their therapy.

REFERENCES

1. Atherton, D. J., Sheehan, M. P., & Rustin, M. Trditional Chinese plants for eczema. *Lancet*, 1991, 338, 510.

2. Cassileth, B. R., Lusk, E. J., Miller, D. S., Brown, L. L., Miller, C. Psychosocial correlates of survival in advanced malignant disease. *New England Journal of Medicine,* 1985, 312, 1551–55.
3. Clavel, F., & Paoletti, C. A study of various smoking cessation programs based on close to 1000 volunteers recruited from the general population: 1-month results. *Revue d'Epidemiologie et de Sante Publique,* 1990, 38, 133–38.
4. Dale, A., & Cornwell, S. The role of lavender oil in relieving perineal discomfort following childbirth: A blind, randomised clinical trial. *Journal of Advanced Nursing,* 1994, 19, 89–96.
5. Davis, P. *Subtle aromatherapy*. Saffron Walden: CW Daniel, 1991.
6. Dundee, J. W., Chestnutt, W. N., Ghaly, R. G., & Lynas, A. G. Traditional Chinese acupuncture: A potentially useful anitemetic? *British Medical Journal,* 1986, 293, 583–84.
7. Dundee, J. W., Ghaly, R. G., Fitzpatrick, K. T., et al. Acupuncture to prevent cisplatin-associated vomiting. *Lancet,* 1987, 1083.
8. Dundee, J. W., Sourial, F. B., Ghaly, R. G., & Bell, P. F. P6 acupressure reduces morning sickness. *Journal of the Royal Society of Medicine,* 1988, 81, 456–57.
9. Dundee, J. W., Yang, J. Prolongation of the antiemetic action of P6 acupuncture by acupressure in patients having cancer chemotherapy. *Journal of the Royal Society of Medicine,* 1990, 83, 360–62.
10. Dundee, J. W., Yang, J, & McMillan, C. Non-invasive stimulation of the P6 (Neiguan) antiemetic acupuncture point in cancer chemotherapy. *Journal of the Royal Society of Medicine,* 1991, 84, 210–12.
11. Echt, D. S., Liebson, P. R., Mitchell, L. B., et al. Mortality and morbidity in patients receiving encainide, flecainide, or placebo: The cardiac Arrhythmia Suppression Trial. *New England Journal of Medicine,* 1991, 324, 781–88.
12. Fisher, P. Research into homeopathic treatment of rheumatological disease: Why and how? *Complementary Medical Research,* 1990, 4, 34–40.
13. Fuller, J. A. Smoking withdrawl and acupuncture. *Medical Journal of Australia,* 1982, 1, 25–29.
14. Gellert, G. A., Maxwell, R. M., & Siegel, B. S. Survival of breast cancer patients receiving adjunctive psychosocial support therapy: A 10-year follow-up study. *Journal of Clinical Oncology,* 1993, 11, 66–69.
15. Koes, B. W., Assendelft, W. J., van der Heijen, G. J. M. G., et al. Spinal manipulation and mobilisation for back and neck pain: A blinded review. *British Medical Journal,* 1991, 303, 1298–303.
16. Koes, B. W., Bouter, L. M., van Mameren, H., et al. Randomised clinical trial of manipulative therapy and physiotherapy for persistent back and neck complaints: Results of one year follow up. *British Medical Journal,* 1992, 304, 601–05.
17. Lamontagne, Y., Lawrence, A., & Gagnon, M. A. Acupuncture for smokers: Lack of long-term effect in a controlled study. *Canadian Medical Association Journal,* 1980, 5, 787–90.
18. Mayaux, M. J., Guihard-Moscato, M. L., Schwartz, D., et al. Controlled clinical trial of homeopathy in postoperative ileus. *Lancet,* 1988, 1, 528–29.
19. McMillan, C. M., & Dundee, J. W. The role of transcutaneous electrical stimulation of Neiguan anti-emetic acupuncture point in controlling sickness after cancer chemotherapy. *Physiotherapy,* 1991, 77, 499–502.
20. Price, S., & Price, L. *Aromatherapy for health professionals*. London: Churchill Livingstone, 1995.
21. Rogers, C. R. *On becoming a person: A therapist's view of psychotherapy*. London: Constable, 1961. Quoted by: Mills S. Conflicting research needs in complementary medicine. *Complementary Medicine Research,* 1986, 1, 40–47.
22. Sacks, H. S., Chalmers, T. C., et al. Sensitivity and specificity of clinical trials: Randomized versus historical controls. *Archives of Internal Medicine,* 1983, 143, 753–55.
23. Scheid, V. Orientalism revisited. *European Journal of Oriental Medicine,* 1993, 1, 23–33.

24. Schultz, K. F., Chalmers, I., Hayes, R. J., & Altman, D. G. Empirical evidence of bias. *Journal of the American Medical Association*, 1995, 273, 408–12.
25. Sheehan, M. P., & Atherton, D. J. A controlled trial of traditional Chinese medicinal plants in widespread non-exudative atopic eczema. *British Journal of Dermatology*, 1992, 126, 179–84.
26. Sheehan, M. P., & Atherton, D. J. One-year follow up of children treated with Chinese medicinal herbs for atopic eczema. *British Journal of Dermatology*, 1994, 130 488–93.
27. Sheehan, M. P., Rustin, M. H., Atherton, D. J., et al. Efficacy of traditional Chinese herbal therapy in adult atopic dermatitis. *Lancet*, 1992, 303, 1298–303.
28. Shipley, M., Berry, H., Broster, G., et al. Controlled trial of homoeopathic treatment of osteoarthritis. *Lancet*, 1983, i, 97–98.
29. Taylor Reilly, D., McSharry, C., Taylor, M. A., & Aitchison, T. Is homeopathy a placebo response? Controlled trial of homeopathic potency, with pollen in hayfever as a model. *Lancet*, 1986, II, 881–85.

Clinton Presidential Records
Digital Records Marker

This is not a presidential record. This is used as an administrative marker by the William J. Clinton Presidential Library Staff.

This marker identifies the place of a tabbed divider. Given our digitization capabilities, we are sometimes unable to adequately scan such dividers. The title from the original document is indicated below.

Divider Title: Research Training

VI

Questions for the Research Training Panel:

What training do conventional investigators need to conduct CAM research?

Are there specific issues involved in training conventional researchers to study CAM questions?

What training do CAM professionals need to (1) collect valid data; (2) collaborate on research projects; (3) contribute as members of research teams; (4) design and conduct high quality research?

How can CAM practitioners become better educated in the critical evaluation of CAM product claims and the results of CAM protocols and clinical research?

What policy recommendations can you offer the Commission on the training of (1) conventional investigators; and (2) CAM investigators and practitioners to strengthen CAM research and the pool of highly qualified researchers in CAM?

Clinton Presidential Records
Digital Records Marker

This is not a presidential record. This is used as an administrative marker by the William J. Clinton Presidential Library Staff.

This marker identifies the place of a tabbed divider. Given our digitization capabilities, we are sometimes unable to adequately scan such dividers. The title from the original document is indicated below.

Pearson/West

Divider Title: _____

Biosketches for the 5/2001 White House Commission

Dr. Nancy J. Pearson received her Ph.D. degree in Biology from Brandeis University and was an NRSA postdoctoral fellow at Albert Einstein College of Medicine (NY). Her area of research was microbial molecular genetics. She was on the faculty of the University of Maryland until 1989, after which she accepted a position at NIH as a Scientific Review Administrator in the newly formed Genome Center. In 1993, she moved to the NIH Center for Scientific Review, eventually becoming Chief of the Genetic Sciences Initial Review Group. Dr. Pearson joined NCCAM as a Program Director in February 2001.

Dr. Neal B. West received his Ph.D. degree in Botany from the University of Chicago and did biochemical research for 20 years at the Oregon Regional Primate Research Center, investigating the mechanisms of steroid hormone action in the reproductive tracts of non-human primates. He came to the NIH in 1990 and trained for one year as a Grants Associate before serving as a Scientific Review Administrator at the NCI for almost 5 years. Dr. West joined the NCRR on August 5, 1996 and served for over four years as the Program Director for Laboratory Animal Sciences in the Comparative Medicine area. He accepted his current position of Program Director at NCCAM in July of 1999.

WHITE HOUSE COMMISSION

Training of Conventional and CAM Investigators
5/16/2001

Drs. Nancy J. Pearson and Neal B. West
National Center for Complementary and Alternative Medicine
National Institutes of Health

The NCCAM's ability to achieve its research goals is dependent on the availability of a critical mass of skilled investigators in both CAM and conventional communities. Thus, NCCAM must encourage skilled researchers to investigate CAM approaches and train CAM and conventional practitioners to conduct or participate in rigorous CAM research studies. To this end, NCCAM is committed to providing funding for high quality mentored research training and career development opportunities to predoctoral and postdoctoral students who are interested in pursuing a career in CAM research. Funding for research training and career development opportunities are available not only to CAM practitioners, so that they may gain the knowledge and experience to conduct rigorous research in their field, but also to conventional medical researchers and practitioners so that they may increase their knowledge and experience in specific CAM research areas. In addition, NCCAM is committed to increasing the number of trainees from under-represented populations in scientific research. To this end, a request for applications (RFA) has been issued and $500,000 has been made available in FY 2001 to fund two high quality NCCAM Institutional Research Training Grants for Minority Researchers. The overall goals of all NCCAM training and career development grant initiatives are to increase the number, quality, and diversity of CAM researchers.

Since the NCCAM wishes to take advantage of all options for research training, our Center has agreed to support the full array of NIH research training grant support mechanisms (especially utilizing the National Research Service Award Program [NRSA]) and career development grant support mechanisms across all possible career stages. These mechanisms are listed on the chart on the next page by stage of training/career development.

NCCAM Research Training and Career Development Awards

Stage of Career — Mechanism of Support

Undergraduate

- Research Supplements for Underrepresented Minorities
- Research Supplements for Individuals with Disabilities

Graduate School

Awards to Individuals

- F31 - NCCAM NRSA Individual Predoctoral Fellowship
- F31 - NRSA Predoctoral Awards to Minority Students
- F31 - NRSA Predoctoral Awards to Students with Disabilities

Awards to Institutions

- T32 - NRSA Institutional Research Training Program
- T32 - NCCAM NRSA Institutional Research Training Program for Minority Researchers
- T35 - NRSA Short-Term Institutional Research Training Grant

Postdoctoral *

Awards to Individuals

- F32 – NCCAM NRSA Postdoctoral Fellowship
- K01 - Mentored Research Scientist Development Award
- K07 - Academic Career Award (Development)
- K08 - Mentored Clinical Scientist Development Award
- K23 - Mentored Patient-Oriented Research Career Development Award

Awards to Institutions

- T32 - NRSA Institutional Research Training Program
- T32 - NCCAM NRSA Institutional Research Training Program for Minority Researchers

Independent Career

Awards to Individuals

- F33 - NRSA Senior Fellowship
- K02 - Independent Scientist Award
- K05 - Senior Scientist Award
- K07 - Academic Career Award (Leadership)
- K24 - Mid-Career Investigator Award in Patient-Oriented Research

* This includes applicants who have received an N.D., D.C., O.M.D., D.O., Ph.D., M.D., D.D.S., D.V.M., O.D., D.P.H., D.Sc., Eng.D., Dr. Ph.D., Pharm.D., D.S.W., D. Psy. or equivalent doctoral degree from an accredited domestic or foreign institution .

Even before 1999, when NCCAM obtained the authority to issue our own awards, we committed funds to support fellows through other NIH Institutes (copays) and other Federal agencies. And we are following through on those commitments already made. In FY 99, our first year of funding, 1.31 percent of NCCAM's budget went to support research training in the form of Individual and Institutional NRSA Training Grants and Mentored Research Career Development Awards. In FY 2000, this increased to 3.78 percent of NCCAM's budget. More detailed information is given in the Table below on the distribution of funds for FY2000.

NCCAM TRAINING & CAREER DEVELOPMENT SUMMARY

- NCCAM Fellowships

  7 F31 (NRSA Predoctoral Fellowships)
  2 F32 (NRSA Postdoctoral Fellowships)
  (subtotal) 9 Fs

- Institutional Training Grants (T32)

  5 Copays (1 pre- and 6 postdoctoral trainees) (3 are chiropractors)
  2 NCCAM's (4 pre- and 8 postdoctoral trainees)

  (subtotal) 7 T32s (5 pre- and 14 postdoctoral trainees)

TOTAL TRAINING (NRSA) (12 pre- and 16 postdoctorals) $1.1 M

- Career Development Awards

  3 K01s
  2 K08s
  4 K23s (1 is a chiropractor)
  2 K24s

TOTAL CDAs (11 postdoctorals) $1.3 M

NCCAM TOTAL FY 2000 TRAINING & CAREER DEVELOPMENT: $2.4 M

In addition to these NRSA research training and career development grant support mechanisms, training in CAM also occurs within several of NCCAM's 12 large Specialty Centers. NCCAM-funded regular research grants also provide training in CAM for students, fellows, technicians, and others who are key personnel on these grants. Finally, in the near future, NCCAM plans to further enhance its ability to support research training in CAM by the establishment of an interdisciplinary, intramural research program on the NIH campus.

Clinton Presidential Records
Digital Records Marker

This is not a presidential record. This is used as an administrative marker by the William J. Clinton Presidential Library Staff.

This marker identifies the place of a tabbed divider. Given our digitization capabilities, we are sometimes unable to adequately scan such dividers. The title from the original document is indicated below.

Blanks

Divider Title: _____

White House Commission on Complementary and Alternative Medicine Policy
Research Challenges: "Training of Conventional and CAM Investigators"

Topic Time: Wednesday May 16, 2001, 9:25-10:05

Speaker: Robert H. Blanks, Ph.D.
Professor of Anatomy and Neurobiology
University of California, Irvine

Outline of testimony:

1. Speakers expertise in research training of Conventional/CAM practitioners
   - Human research compliance: Chair, Institutional Review Board (IRB)
   - Animal research compliance: Chair, Institutional Animal Care and Use committee (IACUC)
   - Extensive research experience in animal and human research
   - Medical educator: 23 years
   - Director of Research for large clinical service
   - Lead Researcher: practice-based outcomes in several CAM modalities
   - Director of multidisciplinary CAM research group
   - Area of CAM research: subluxation-based chiropractic, Oriental Medicine)
   - Research expertise: health, wellness and quality of life outcomes assessment

2. Summary of research strategies in medical model vs. social science model of research

   Although the medical model and its research (gold) standard is the randomized clinical trial, there are other well-established research formats in Social Science which are ideally suited for creating a "culture of evidence" around the broad, holistic perspectives of many of the CAM modalities. A brief discussion of these often conflicting research designs (medical vs. social science models) should help set the stage for a systematic approach for evaluating the CAM modalities which, by their very nature, respond well to social science survey research, but do not lend themselves to "sham interventions", "randomization" or other criteria of the medical model.

3. Recommended focus of CAM research agenda:

Research design:	Focus on practice-based health outcomes studies (initial cross-sectional studies, then time-varying series and selected randomized, controlled studies) Care to evaluate CAM modalities by discipline, and in many instances subdiscipline to avoid the experimental errors from pooled data from dissimilar clinical protocols.
Outcomes measures:	broad measures of health (bio-, psycho-, social-, & spiritual-measures) vs. symptom relief
Research populations:	Community-based population studies to control for sociodemographic, socioeconomics, and confounding health issues (e.g., health status, social networks, mental/emotional status, health beliefs/values, life-style practices, and concurrent use of other health promoting modalities.
Research Methods:	Multivariate analysis to evaluate multiple outcome measures

4. Research training formats:
   - Regional training and statistical core sites funded by core center grants.
   - Coordinated training sessions: National conferences, Continuing Education, etc.
   - Coordinated CAM research training curriculum: specialized test, visuals, multimedia format
   - Degree and non-degree programs (Specialized MS or Certificate programs for combined training (e.g., MD-MS, DO-MS, DC-Ph.D., etc.)

Withdrawal/Redaction Marker
Clinton Library

DOCUMENT NO. AND TYPE	SUBJECT/TITLE	DATE	RESTRICTION
001. resume	Robert Blanks (partial) (1 page)	nd	b(6)

COLLECTION:
Federal Records
White House Commission on Complementary and Alternative Medicine Policy [WHCCAMP]

OA/Box Number: 43232

FOLDER TITLE:
WHCCAM Commissioner Binder, Day 3, May 16, 2001 [2]

2011-0568-S
kc813

RESTRICTION CODES

Presidential Records Act - [44 U.S.C. 2204(a)]

P1 National Security Classified Information [(a)(1) of the PRA]
P2 Relating to the appointment to Federal office [(a)(2) of the PRA]
P3 Release would violate a Federal statute [(a)(3) of the PRA]
P4 Release would disclose trade secrets or confidential commercial or financial information [(a)(4) of the PRA]
P5 Release would disclose confidential advice between the President and his advisors, or between such advisors [a)(5) of the PRA]
P6 Release would constitute a clearly unwarranted invasion of personal privacy [(a)(6) of the PRA]

C. Closed in accordance with restrictions contained in donor's deed of gift.
PRM. Personal record misfile defined in accordance with 44 U.S.C. 2201(3).
RR. Document will be reviewed upon request.

Freedom of Information Act - [5 U.S.C. 552(b)]

b(1) National security classified information [(b)(1) of the FOIA]
b(2) Release would disclose internal personnel rules and practices of an agency [(b)(2) of the FOIA]
b(3) Release would violate a Federal statute [(b)(3) of the FOIA]
b(4) Release would disclose trade secrets or confidential or financial information [(b)(4) of the FOIA]
b(6) Release would constitute a clearly unwarranted invasion of personal privacy [(b)(6) of the FOIA]
b(7) Release would disclose information compiled for law enforcement purposes [(b)(7) of the FOIA]
b(8) Release would disclose information concerning the regulation of financial institutions [(b)(8) of the FOIA]
b(9) Release would disclose geological or geophysical information concerning wells [(b)(9) of the FOIA]

[001]

PROFESSIONAL RESUME — DR. ROBERT H.I. BLANKS

Office: (949) 824-7991
Department of Anatomy and Neurobiology
University of California, Irvine
Irvine, California 92697-1275 USA
FAX: (949) 824-8549
e-mail: rhblanks@uci.edu

(b)(6)
Rochester Hills, Mi 48306

EDUCATION

Ph.D. Physiology, University of California, Los Angeles, CA 1973
M.A. Physiology, University of California, Los Angeles, CA 1969
B.A. Pre-Medical Sciences, Humboldt University, Eureka, CA 1968

EMPLOYMENT

Tenure- Track Appointment, Dept Anat & Neurobiol, UC Irvine, Irvine, CA
Professor, 1993 – present
Associate Professor, 1983 – 1993
Assistant Professor, 1978 - 1983
Joint Appointments, UC Irvine, Irvine, CA
Dept Otolaryngol Head/Neck Surg 1978 – 1998
School of Biological Sciences 1980 - present
Dept. Sociology 1999 - present
Visiting Prof, Anatomy, South-Baylo, Oriental Med Univ, Anaheim, CA, 7/96- present
Associate in Anatomy, Dept Anatomy, Harvard Medical School, 1977 - 1978
Visiting Scientist, Max Planck Inst. Brain Res, Frankfurt, Germany 1974 - 1976
Asst Res Neurologist, Depart Neurol, UC Los Angeles, 1973 - 1974
Pre-Doctoral Fellowship, Depart Physiol, UC Los Angeles, 1969 - 1973

AFFILIATIONS

Clinical Anatomists Association
Council of Chiropractic Practice, Chandler, AZ
International Oriental Medical Research Association, Irvine, CA
Association for Assessment and Accreditation of Laboratory Animal Care (AAALAC)
(consultant and site visitor)

ADMINISTRATION

Human Subjects Research Compliance (1999 – present):
(Committee member and Chair UCI Institutional Review Board (IRB): Human subjects (IRB) compliance, review of human subjects research protocols and research design)

Animal Research Oversight and Compliance (1981–1991):
(Chairman of Institutional Animal Care and Use Committee (IACUC); Monitored compliance with Federal (USDA, NIH-OPRR), State and University Regulatory issues; Mentored new committee members, faculty and students on guidelines and regulatory issues).

Animal Research Administration (1991 – 1993):
(Advisor to Vice Chancellor for Research and Graduate Studies on all animal research conduct and regulations issues)
Faculty Asst. Vice Chancellor for Res & Grad studies 1991 – 1993
Interim Director, UCI Office of Animal Research 1992 – 1993
Interim Animal Research Facilities Manager 1992
Chairman, Campus Veterinarian Search Committee 1993

ADMINISTRATION (cont.)

Chairman, Scientific Misconduct Panel (1997 – 1999):
(Provided oversight to one of the largest scientific misconduct panels ever convened at the University of California. The case involved alleged misconduct across 10 years by multiple investigators, and involves millions of dollars in class action lawsuits from hundreds of respondents).

Co-developer and Co-Director (1995–1997): for a multi-institutional, multidisciplinary Southern California Center for the Study of Complementary Medicine. This concept established a solid working affiliation between the UCI College of Medicine, a major chiropractic college, and one of the largest Schools of Oriental medicine in Southern California.

Director of Medical Education, Dept of Anat & Neurobiol, UCI (1994–1998):
(Coordinated curriculum and scheduling for the three departmental courses – gross anatomy/embryology, neuroscience, and histology. Was successful in implementing two new courses into the medical curriculum 1) a 6-week intensive summer program for remediation in gross anatomy, and 2) "Introduction to Oriental Medicine" an interdisciplinary course for freshman medical students to teach cross-cultural, and complementary medical care).

Chairman, Academic Personnel Committee, Dept of Otolaryngol (1987–1997):
(In change of committee responsible for all faculty appointments, and preparation of merit/promotional packets for 20 member clinical faculty).

RESEARCH

Distinguished (25 year) basic science and clinical research career:
(specializing in visual-vestibular interaction, postural disorders, sleep disorders, CAM health outcomes research: health, wellness and quality of life assessment).

Directed the research activities of the following students (career total):

11	Graduate students and postdoctoral fellows
78	Undergraduate biological science honors students
19	Surgery and Otolaryngology, Head/Neck Residents
10	Medical students
5	Physical Therapy Masters students

Director of Research, Dept. Otolaryng. Head/Neck Surgery (1981-1989):
(Supervised resident research; mentored faculty research development)

Research Career Development Award (1979 – 1984):
(NIH – National Eye Institute)

Principal Investigator on International, CAM health outcomes research (1994 – present):
(practice-based; retrospective and prospective design; patient-centered study on health values and beliefs using wellness and quality of life surveys).

PUBLICATIONS (Selected from 65 peer-reviewed manuscripts, 12 book chapters and 88 abstracts)

Blanks, R.H.I., Fowler, C.G., Zizz, C.A., Williams, K.E. Postural adjustments by moving visual (horizontal optokinetic) patterns. J. Am. Acad. Audiol. 7:39-48, 1996.

Blanks, R.H.I., Schuster, T. L., Dobson, M. A retrospective assessment of network care using a survey of self-rated health, wellness and quality of life. J. Vertebral Subluxation Res. 1:15-31, 1997.

Blanks, R.H.I., Giolli, R.A., and van der Want JJL. Neuronal Circuitry and Neurotransmitters in the Pretectal and Accessory Optic Systems. In: Progress in Brain Research, Neurochemistry of the Vestibular System, A.J. Beitz and J.H. Anderson (eds.), CRC Press, Boca Raton, FL, 1999.

EXTRAMURAL FUNDING (Career summary)

International	NATO Research Travel Grant	22,704
	Korean Manpower Institute	20,000
	Migun Medical Instruments, Korea	551,781
Federal	National Institutes of Health	1,024,273
	National Science Foundation	349,615
	Veterans Administration	752,352
	National Aeronautics & Space Administration	663,510
State	University of California	23,306
Private	Respiratory Sleep Disorders Foundation of CA.	250,000
	Pauley Research Foundation.....	20,000
	Margaret W. & Herbert Hoover, Jr. Foundation	49,775
	Forty First Medical Endowment Trust	32,000
	Association for Network Chiropractic	244,000
	Wellness Research Institute, Body Wise Int.	30,000
	Career (Total Direct Costs)	4,033,316

TEACHING

Medical and allied health education and curriculum development (1978-present):
(Teaching specialties: gross anatomy, neuroscience, and alternative/complementary medicine approaches to stress reduction)

Lecturer, Gross Anatomy, Neuroscience	1979 - present
Course Director, Medical Gross Anatomy	1988 – present
Course Director, Summer Gross Anatomy	1995 – present
Course Director, Introduction Oriental Med	1998 – present
Course Director, "Acupuncture Anatomy"	1994 - present
Continuing Education in Acupuncture	1999 - present

Taskforce Member, Development of Ph.D. program in Audiology (1989-1992):
(Assisted in developed curriculum and all affiliations for a new graduate program Audiology involving three Institutions, and several Departments at the UCI College of Medicine).

HONORS

Research Career Development Award, National Eye Institute	1979-1984
Silver Beaker Certificate for all around basic science faculty	1982, 1983
Kaiser-Permanente Award for excellence in teaching	1994
Vice Chancellor's Award for committed service to Animal Programs	1994
Award of Appreciation, South Baylo Univ, curriculum development	1995
Award for Extraordinary Service, UCI, Scientific Misconduct panel	1997
Excellence in Teaching, South Baylo Univ School of Oriental Med	1998

Clinton Presidential Records
Digital Records Marker

This is not a presidential record. This is used as an administrative marker by the William J. Clinton Presidential Library Staff.

This marker identifies the place of a tabbed divider. Given our digitization capabilities, we are sometimes unable to adequately scan such dividers. The title from the original document is indicated below.

Phillips

Divider Title: _____

Dr. Russell Phillips is Associate Professor of Medicine and Director of Research in the Division of General Medicine at Beth Israel Deaconess Medical Center and Harvard Medical School. He directs Harvard-wide fellowship programs in complementary and alternative medicine, and general medicine. His research interests include clinical decision-making and epidemiology, and CAM research. He directed Beth Israel's participation in the SUPPORT Project, a study of patient outcomes, preferences and decision-making in the care of seriously ill patients. His work includes studies of patient preferences for treatment and physician-patient communication, treatment effectiveness and disease outcomes and the impact of patient age and race on resource use and clinical decision-making. He is Principal Investigator of a study on quality of care for Asian-Americans funded by ARHQ and is co-investigator on an NCCAM-funded project comparing usual care to a choice of massage, acupuncture, or chiropractic for patients with low back pain. He is Principal Investigator on a proposal to study the effects of massage on patients hospitalized with cancer. Dr. Phillips is also Principal Investigator on an NCCAM funded T32 training grant, which prepares physicians for academic careers focusing on CAM research. Dr. Phillips has mentored more than 20 fellows. His trainees' publications include alternative medicine use by patients with HIV and the use of alternative medicine by older patients. Current CAM fellows are conducting clinical trials on magnets for patients with osteoarthritis, Tai Chi for patients with heart failure, and chiropractic for children with recurrent ear infections. Dr. Phillips intends to become a leading CAM researcher and mentor, and is applying for an NCCAM funded Mid-Career Investigator Award in CAM research.

Testimony prepared by Russell S. Phillips, M.D. for the White House Commission on Complementary and Alternative Medicine Policy

What training do conventional investigators need to conduct CAM research?

1) Investigators need to be well versed in research methodology, usually requiring advanced training at a Master's or doctoral level.
2) Investigators need an understanding of CAM modalities they wish to study, including the paradigms of health and healing held by CAM practitioners, mechanisms of action, (when these are understood), and familiarity with previous research on the CAM modality of interest.
3) Investigators need mentoring, both by CAM practitioners and experienced clinical investigators, to assist in developing the research question, methods, the analysis plan, and preparation of reports for publication.

At Harvard Medical School, we have a training program for internal medicine physicians to develop skills needed to perform CAM research. This program provides Master's level training in research methods, introduction to CAM modalities and practitioners, and mentoring, both by CAM practitioners and experienced clinical investigators.

I would like to describe our program briefly:

The Faculty Development and Fellowship Program in Complementary and Alternative Medicine at Harvard Medical School accepts applicants for 3-year academic fellowships in General Internal Medicine, and Complementary and Alternative Medicine. The Program offers each Fellow an appointment at Harvard Medical School hospitals. All fellows participate in the intensive summer program in Clinical Effectiveness at the Harvard School of Public Health. Fellows qualifying for acceptance to the Harvard School of Public Health pursue a rigorous curriculum that could lead to a Master of Science or Master of Public Health degree. The Program also includes structured experiences to improve teaching skills and supervised clinical activities under the direction of experienced faculty in general internal medicine, and complementary and alternative medicine. Each Fellow is expected to design, conduct, present, and publish one or more original investigative projects. This fellowship in Complementary and Alternative Medicine is funded by the National Institutes of Health, National Center for Complementary and Alternative Medicine, as a T32 training program.

Our program begins each July 1 with a seven-week summer core curriculum taught by Program faculty in collaboration with faculty at the Harvard School of Public Health (Clinical Effectiveness Program). The 15-credit core curriculum includes required courses in biostatistics, epidemiology, and elective courses in health policy, health services research, decision sciences, quality improvement, and public health ethics. The intensive summer experience includes about six hours per day of classroom time and about four hours per night of assignments, including two required class presentations.

After completion of the summer core curriculum, Fellows accepted into the Master's Program may continue to take advanced courses at the Harvard School of Public Health and, upon earning 40 academic credits, can obtain a Master of Science or Master of Public Health degree. Although the degree itself is optional, Fellows are strongly encouraged to take specific advanced course work that will help develop the investigative skills required for their original research projects.

Fellows attend weekly seminars or research conferences addressing research in progress, and study design. These sessions, which bring together Fellows who are located at various sites within the program, reinforce the skills learned in the classroom. Monthly seminars in the Center for Alternative Medicine Research and Education at Beth Israel Deaconess Medical Center provided a forum to present and discuss ongoing research, and a journal club provides opportunities to review recently published research in complementary and alternative medicine.

Beginning at the end of the summer core curriculum of the first fellowship year, each Fellow provides care for patients. Fellows also works with alternative medicine providers, observing their practices. An integrative practice site at Beth Israel Deaconess Medical Center is under development, and will likely serve as a clinical site in the future.

Each Fellow participates in a series of "teaching how to teach" seminars and colloquia under the direction of clinicians who are recognized as outstanding teachers. In addition, each fellow has opportunities to serve as a teacher of medical students and housestaff in the ambulatory practice of a sponsoring clinical site, and to teach undergraduate courses and continuing medical education programs in alternative medicine at Harvard Medical School.

Each Fellow works in close collaboration with a faculty mentor and initiates a research project during the first year of the fellowship. The project may use any or all of the methodologic disciplines that are included in the academic curriculum and focuses on complementary and alternative medicine. Projects may utilize a variety of methodologies and topics that include: clinical epidemiology, clinical trials, decision analysis, patient outcomes, health services research, health policy, practice variation, technology assessment, access and equity in health care, quality improvement, disease prevention, health promotion, ethics, and informatics. Fellows present their plans and research in progress at seminars within the program. They are encouraged and expected to present their findings at regional and national meetings of investigators, such as the upcoming International Scientific Meeting on CAM research in San Francisco. They will also be invited to present their work at annual Continuing Medical Education courses directed by the Center for Alternative Medicine Research and Education at Beth Israel Deaconess Medical Center.

Faculty in the Program are recognized leaders in academic general internal medicine and complementary and alternative medicine, and include experienced CAM researchers and practitioners. Members of the core faculty include: Drs. Joseph Audette (acupuncture), Edward Chapman (homeopathy), E. Francis Cook (epidemiology), Roger Davis

(biostatistics), David Eisenberg (CAM research), Robert Fletcher (vitamins and disease prevention), Peter Goldman (clinical pharmacology), Janet Kahn (massage), Ted Kaptchuk (oriental medicine), Eric Leskowitz (hypnosis), Weidong Lu (acupuncture), and Russell Phillips. Dr. Phillips directs the program, and Dr. Eisenberg, Director of the Center for Alternative Medicine Research and Education, serves as Co-director.

Applicants must be board-eligible or board-certified in internal medicine at the start of their first fellowship year. Applicants are screened on the basis of their career goals and on the recommendations of faculty from their medical schools and residency programs.

Fellows who are accepted into the Program are then required to complete a formal application to the Harvard School of Public Health prior to beginning the Fellowship. In addition to the fellowship stipend, the Program pays full tuition for the summer core curriculum, and for fellows in the Master's degree program, all the courses that are required (40 credits) for a Master's level degree at the Harvard School of Public Health. The sponsoring Harvard institution pays the cost of tuition not covered by our NIH training grant.

The program is particularly interested in applications from individuals from under-represented minority groups.

Are there specific issues involved in training conventional researchers to study CAM questions?

Conventional researchers need to understand the CAM practice being evaluated, and need to work collaboratively with CAM practitioners who are expert in the CAM modality being studied. Several research issues are potentially problematic in performing CAM research. These include consideration of the placebo effect, difficulties in blinding both investigators, and difficulties standardizing interventions. By its nature, many CAM treatments, such as massage, are individualized for the patient based on the provider's expertise. Not allowing the provider to individualize the therapy provided may reduce substantially the effects of the intervention under study and would not reflect CAM practice in the community.

What training do CAM professionals need to (1) collect valid data; (2) collaborate on research projects; (3) contribute as members of research teams; (4) design and conduct high quality research?

These research activities are listed in order of increasing complexity. For each of these activities, the necessary training would be similar to that required of other clinicians who participate in research. Commonly clinicians can be trained, in relatively short training sessions, to collect and record data. Members of our research group did this successfully in a NCCAM-funded study of patients seen by chiropractors, acupuncturists, naturopaths and massage therapists. Training sessions should include the fundamentals of research methods, how to collect and store scientific data, and human subject protections that need

to be in place before research can be performed. For CAM providers to collaborate and contribute as members of a research team, CAM providers need increasing research sophistication which may vary depending on the study under consideration. In order to design and conduct high-quality research, CAM providers need the same training we provide to physician-researchers, as described above.

How can CAM practitioners become better educated in the critical evaluation of CAM product claims and the results of CAM protocols and clinical research?

Similar to the training physicians should receive, CAM practitioners need training in skills necessary to critically appraise a published study. These skills include a basic understanding of research design and the limits of each approach, an understanding of different measures of effect, of the meaning of p values, statistical significance and confidence intervals, and potential sources of bias and confounding. An excellent series of papers on critical appraisal skills has been published in JAMA, which form the basis for many critical appraisal courses taught to physicians, and could easily be adapted for CAM practitioners. Ideally, courses on critical appraisal skills should be taught as part of continuing education and undergraduate programs. Given the limited regulation of CAM product claims, it is especially important for CAM providers to be able to assess the strength of evidence supporting these claims.

What policy recommendations can you offer the Commission on the training of (1) conventional investigators; and (2) CAM investigators and practitioners to strengthen CAM research and the pool of highly qualified researchers in CAM?

First, I recommend that increased funding for CAM research be made available for training programs for both conventional investigators and CAM investigators and practitioners. We need increased funding for career development awards, to enable investigators focusing on CAM to develop into independent investigators. We need increased funding for mid-career investigators to provide support for the time required to mentor new investigators, and we need increased research support for CAM, to further stimulate the growth of CAM research.

TWO "WORLD" VIEWS OF RESEARCH & RESEARCH TRAINING

	"BIOMEDICINE MODEL"	"SOCIAL SCIENCE MODEL"
RESEARCH OBJECTIVES:	seeks laws, studies objects prefers simplification linear logic distrustful of peoples opinions collects facts limits choices assumes data are objective	studies patterns, processes & contectual format skeptical of simplification search for patterns assumes peoples knowledge matters track rivers of information evaluates many choices assume data are subjective
CLINICAL OBJECTIVES:	reductionism disease care benchmarks: minimum standards patient as (passive) recipient of care Third-party (Insurance) reimbursed	holism/vitalism illness/wellness promotion benchmarks: vital standards client-practitioners health partnerships Insurance reimbursed short-term; often self-paid thereafter
EVIDENCE-BASED (PRACTICE) GUIDELINES:		
Chiropractic:	Musculoskeletal Chiropractic (Guidelines for Chiropractic Quality Assurance and Practice Parameters: Proceedings of the Mercy Center Consensus Conference, 1992)	Subluxation-based Chiropractic (Council of Chiropractic Practice Guidelines, Clinical Practice Guideline No 1: Vertebral Subluxation in Chiropractic Practice, 1998)

Submitted by: Robert H. Blanks, Prof. Anatomy and Neurobiology, University of California, Irvine

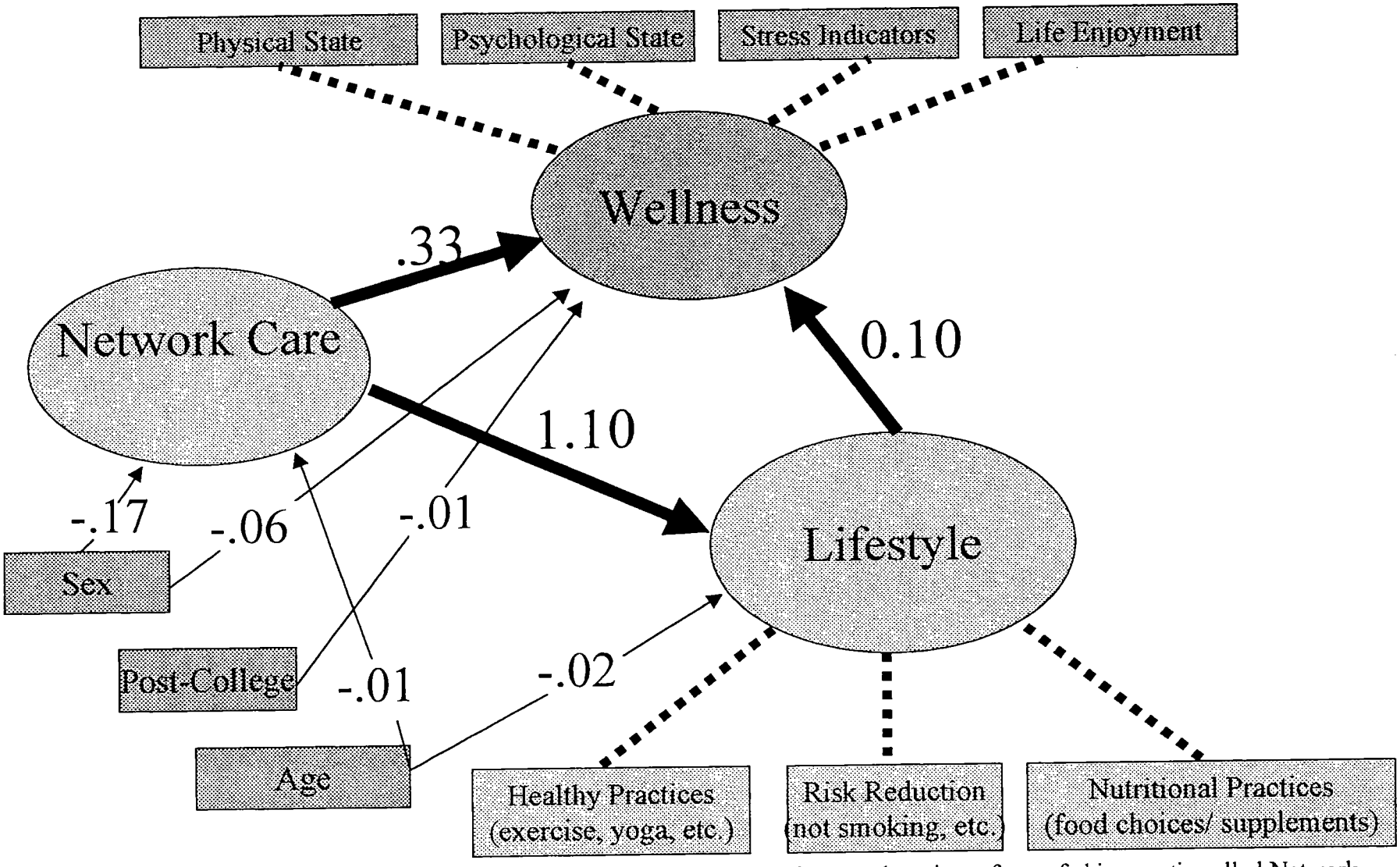

Figure 1: "Wellness" model derived from patient-centered data from 2,818 patients undergoing a form of chiropractic called Network Spinal Analysis (NSA). The cause-effect relationships between the "derived variables" in the ovals (NSA Care, Wellness, Life Changes, Lifestyle) and the actual items on the questionnaire in rectangles (independent variables) is determined statistically using Path Analysis (for additional details see text). The solid arrows indicate the ordered, cause-effect relationships predicted by the data; the dashed lines indicate the actual variables significantly related to each derived variable. Note that care impacts wellness directly, and indirectly by promoting beneficial lifestyle changes. In this model, lifestyle promotes wellness only indirectly by impacting life enjoyment.

ORIGINAL RESEARCH

A Retrospective Assessment of Network Care Using a Survey of Self-Rated Health, Wellness and Quality of Life.

Robert H. I. Blanks, Ph.D,[1,2]
Tonya L. Schuster, Ph.D,[3] *Marnie Dobson, B.A.*[1]

Abstract — The present study represents a retrospective characterization of Network Care, a health care discipline within the subluxation-based chiropractic model. Data were obtained from 156 Network offices (49% practitioner participation rate) in the United States, Canada, Australia, and Puerto Rico. Sociodemographic characterization of 2818 respondents, representing a 67-71% response rate, revealed a population predominately white, female, well-educated, professional, or white collar workers. A second objective of the study included the development and initial validation of a new health survey instrument. The instrument was specifically designed to assess wellness through patients' self-rating different health domains and overall quality of life at two "time" points: "presently" and retrospectively, recalling their status before initiating care ("before Network"). Statistical evaluation employing Chronbach's alpha and theta coefficients derived from principle components factor analyses, indicated a high level of internal reliability in regard to the survey instrument, as well as stable reliability of the retrospective recall method of self-rated perceptions of change as a function of duration of care. Results indicated that patients reported significant, positive perceived change ($p < 0.000$) in all four domains of health, as well as overall quality of life. Effect sizes for these difference scores were all large (>0.9). Wellness was assessed by summing the scores for the four health domains into a *combined wellness scale,* and comparing this combined scale "presently" and "before Network." The difference, or "wellness coefficient" spanning a range of -1 to +1, with zero representing no change, showed positive, progressive increases over the duration of care intervals ranging from 1-3 months to over three years. The evidence of improved health in the four domains (physical state, mental/emotional state, stress evaluation, life enjoyment), overall quality of life from a standardized index, and the "wellness coefficient," suggests that Network Care is associated with significant benefits. These benefits are evident from as early as 1-3 months under care, and appear to show continuing clinical improvements in the duration of care intervals studied, with no indication of a maximum clinical benefit. These findings are being further evaluated through longitudinal studies of current populations under care in combination with investigation of the neurophysiological mechanisms underlying its effects.

Key Words: Network spinal analysis, vertebral subluxation, chiropractic, self-rated outcomes assessment, wellness, overall quality of life.

Introduction

Network Care is a health care discipline within the subluxation-based chiropractic model[1] practiced by members of the Association for Network Chiropractic (ANC), nationally and internationally. Building from a base of consistent clinical observations, and repeated anecdotal reports of health benefits, the present study was conducted to fulfill the following three objectives: (1) to characterize the patient population undergoing Network Care; (2) to develop a new survey instrument of sufficient design and scope to allow assessment of a non-medical health discipline, and (3) to assess changes in patients' self-rated health, wellness, and overall quality of life.

Network Care is founded on the premise that individuals free of the complex of factors precipitating from, or leading to, vertebral subluxation experience a greater range of inherent adaptability and, hence, a greater sense of relative health or wellness. In a large percentage of individuals, Network Care evokes spontaneous self-perpetuating contractions of the paraspinal musculature.[1] The movements may be subtle, barely perceptible, or very

[1]Department of Anatomy and Neurobiology, [2]Department of Otolaryngology-Head and Neck Surgery, College of Medicine, [3]Department of Sociology, University of California Irvine, Irvine, CA 92697.

Address reprint requests to:
Robert H. I. Blanks, Ph.D., Professor, Department of Anatomy and Neurobiology, College of Medicine, University of California, Irvine, California 92697-1275.
e-mail: rhblanks@uci.edu

obvious and may involve the arms and legs. Over a period of several months, physiological and psychological changes have been observed clinically, including increased flexibility of the spine, increased range of motion, improved mood and sense of relaxation, self-reported "wellness," and greater capacity to cope with stressful situations. These observations provide a basis for considering that Network Care involves body-mind interactions.

Consistent with the definition of health offered in 1958 by the World Health Organization,[2] as being "a state of complete physical, mental, and social well-being, and not merely the absence of disease or infirmity," overall improved health is one of the clinical objectives of Network Care. Until recently, methods to assess health changes relative to this definition have been unavailable.

However, with the advent of new outcome measurements for physical, mental and social well-being, as well as a greater awareness of the need to provide patient-centered outcomes, many disease treatment modalities are beginning to evaluate the effectiveness of their intervention in regard to the more holistic WHO definition of health.[3-4] Ironically, while this adds a greater understanding of "holistic" health to the medical perspective, this approach is not yet widely adopted in the area of health care where it would be most applicable, i.e., non-medical practices that have as their primary clinical goal the enhancement of overall health. This is perhaps due to the disease-specific orientation found in most recently developed surveys.[5]

A more thorough investigation of the efficacy of non-medical approaches is also confounded by a confusion of terms arising from the changing medical perspective. This new lexicon developed to accommodate the increased public and scientific interest and utilization of various practices considered outside of orthodox (allopathic) medicine. Terms such as "complementary" and "alternative" medicine, while useful in categorizing practices which treat disease but are outside of orthodox medicine (e.g., homeopathy, acupuncture), fail to differentiate those practices (non-medical) whose clinical objectives do not involve direct intervention in symptom alleviation or the treatment of disease.

By grouping all health care disciplines considered "outside" orthodox medicine under the umbrella of "complementary/alternative medicine," and then requiring that acceptance of these modalities be preceded by an evidence base derived from randomized clinical trials substantiating disease-treatment efficacy, the non-medical objective of certain disciplines (e.g., subluxation-based chiropractic) is ignored. For example, a specific practice may not be effective in improving the symptoms of arthritis, although it may be effective in improving other aspects of health such as physical, mental, or social well being. Consequently, dismissal of a given practice based on its lack of medical efficacy, without appropriate attention to its non-medical effects, does little to identify why it is being utilized or how it is important in the evolving holistic concept of health.

An example can be seen within chiropractic, where one school of thought views musculoskeletal manipulation as a means of treating certain diseases and/or dysfunctions.[6] Chiropractic practiced under this objective would qualify as a drugless, but palliative (symptom-based) therapy. Consequently, it could be argued that its allopathic objective renders it a form of complementary or alternative medicine and thus could be readily assessed using research methodologies commonly employed in medicine. In contrast, another school of thought within chiropractic (to which Network Care aligns) views correction of the condition of vertebral subluxation as a means of enhancing the body's inherent adaptive abilities, thus improving overall health.[6] This objective, identified with chiropractic since its inception[7] and recently re-emphasized by the Association of Chiropractic Colleges,[8] differentiates it from the clinical objectives of medicine. Consequently, it seems imperative that more appropriate methods be developed to assess the outcomes of such non-medical modes of health care.

Studies now suggest that global self-ratings of health, as well as being highly predictive of such fundamental outcomes as mortality or longevity,[9-10] are also an effective and justifiable means of measuring broader health outcomes (i.e., not just the presence or absence of disease or symptoms).[4] However, the psychometric survey instruments currently available were developed predominantly to assess general health and quality of life relative to medical interventions,[5] and were found lacking for the purposes of the present study. For example, among the 21 most relevant instruments reviewed and evaluated by McDowell and Newell,[5] none are appropriate to measure changes in a population of patients likely to be presenting for care already in relatively good health (an assumption based on numerous informal interviews with Network Care patients and practitioners). This was an important consideration in the present study as problems exist with "ceiling" effects in many of the instruments reviewed which limit the ability to measure improvement in an already healthy population. Additionally, their lack of breadth or excessive detail,[5] led to the development of a new instrument that would more directly reflect the WHO domains of physical, mental, and social well-being, without being cumbersome.

As with any new survey instrument, the first step involved tailoring the design of the items and scales for the particular condition (wellness) being tested, and beginning the validation process by examining the ability of the instrument to detect change (i.e., responsiveness) in a particular setting.[11] The subscales of the wellness questionnaire items were formulated to reflect aspects of the broad WHO definition of health; including the domains of physical and mental/emotional state, and intra/inter-personal life enjoyment indicative of physical, mental and social psychological well-being. In addition, the wellness domain termed "stress evaluation" (ability to cope with the demands of the environment), reflected aspects of functional ability incorporated into broader contemporary definitions of health. These domains, as well as the combined wellness scale, were compared to a standardized overall quality of life instrument[12] to assist in the process of validating the new instrument.

Typically, changes in outcome parameters are assessed with an "over-time" study design to meet a requirement for establishing causality (the cause must precede the outcome in time). However, longitudinal data (multiple measurements for the same individual over time) is costly and time-consuming, and the design is predicated on knowledge of the relevant time-frame. Cross-sectional studies of large populations are the appropriate first stage, and provide the opportunity to examine other ran-

domly distributed factors which may influence the outcomes. Without longitudinal data, suggestive evidence of change (and causality) must rely on the retrospective recall of respondents. The use of measurements of this form is limited, since merely asking respondents for current health ratings is ambiguous regarding the necessary cognitive frame of reference (e.g., compared to who/when). Moreover, merely asking for degree of change or improvement in health does not provide information on current levels of wellness (or the levels most common among individuals entering Network Care). Thus, the method of asking respondents to rate their level of health and wellness both "before Network" and "presently" was developed, with the difference between these two scores representing "perceived change." This method provided a richer and more explicitly circumscribed format of response, allowing the respondents to clearly recognize an attempt was being made to elicit their perceived degree of change in wellness.

The new instrument was considered, overall, as a *"wellness survey,"* yielding a *"wellness coefficient"* (reflecting perceived change) derived from the retrospective assessment of the *combined wellness scale.* The instrument was also designed to derive sufficient data to indicate the extent of its internal and external validity for continued use in evaluating Network Care, as well as for potential use in assessing medical and other non-medical health care approaches.

Materials and Methods

Practitioner population

Network Care is currently practiced by licensed chiropractors recognized through their membership in the Association for Network Chiropractic (ANC), an international organization. In this first stage of sampling, the membership of the ANC was designated as the pool through which the patients were surveyed. All offices were solicited regardless of size, and regardless of full- or part-time practice. Because of the number of ANC practitioners, the census was conducted in two parts. The first survey-questionnaires were distributed in November, 1994 to 145 ANC practitioners with a request to have them completed by their patients and returned by December 30, 1994. The second survey distributed identical questionnaires to the remaining 185 ANC practitioners in March, 1995 with a return request of April 30, 1995. A telephone "tree" was established following each mailing to contact offices at least once, to encourage participation, answer questions, and determine reasons for non-participation.

From the 330 registered practitioners in the United States, Canada, Puerto Rico, and Australia, telephone follow-up revealed 9 were not practicing, leaving a total of 321 eligible to participate. From these, responses were obtained from 156 offices, yielding a practitioner participation rate of 49%.

Patient population

In the patient sampling stage, respondents were solicited through these 156 practices according to the following inclusion criteria: 1) under care for 1 month or more as of the survey date, and 2) 18 years of age or older. University of California, Irvine institutional human subject approval was obtained for the study which required respondents to provide verbal consent to participate in the study. Practitioners were also instructed to sample accordingly: "all practice members [patients] under Network Care for longer than one month [as of the survey date] are to be included in this study. If there are more than one hundred (100) practice members in your practice, and it is impossible to include them all, then simply include all those who are in the practice on a particular day or days (if possible, include at least 80 members)." The practitioners returned the completed questionnaires which were entered into a database for analysis.

Response rate

Survey research literature indicates that studies of national scope designed for multiple subgroup comparisons (e.g., by gender, age, duration of care, etc.) typically require sample sizes of 1500-2500.[13] A total of 2,818 completed questionnaires were received from these 156 practitioners. In a community study of this size it is difficult to obtain an accurate rate of return because there is no central registry of patients undergoing care. Thus, one purpose of this study was a first attempt to estimate the size of this population. This was done as follows: *First,* the average size of Network practices was estimated with data from 77 participating ANC doctors completing a question on a "Doctor's Survey" regarding total number of patients in their practice. This information was supplemented with an additional survey of 59 ANC practitioners attending a regional ANC training session. Both of these samples reported a median practice size of 40. Within the 156 participating offices, the total number of patients under care was thus estimated to be 6,240. *Next,* this total was corrected to estimate the sub-sample meeting the inclusion criteria by age and duration of care. This was accomplished with data from a sub-sample of thirty randomly selected offices, which provided sociodemographic information about total patient population. An estimated 5-7% of patients were ineligible because they were younger than 18 years, thus excluding 312-427 patients from the total patient population. Data on duration of care across all patients indicated that 15-17% of the total population was ineligible because they were under care less than one month at the time of the survey, excluding 936-1061 patients. *Finally,* the size of the eligible population was adjusted for sampling techniques employed by the larger practices. Since an estimated 12.4% of offices included more than one hundred practice members, the selection process excluded a projected 771 otherwise eligible patients from the pool. Within these parameters, the estimated range of the eligible population of patients from the 156 participating offices is 3972-4221, corresponding to an estimated response rate of 67-71%.

Procedures and measures

In this cross-sectional study, participants completed a one-time self-report questionnaire, consisting of a total of 87 questions (some containing sub-questions). The first objective, to characterize the patient population, was addressed through thirty-two questions eliciting information concerning: 1) patient sociodemographic characteristics [13 questions], 2) general health and health care history [8 new questions], 3) chiropractic care history [6 new questions], and 4) physical responses to Network Care [5 new questions]. The second objective of this

study was the development and evaluation of an appropriate outcome instrument. Therefore, patients were also asked to complete 55 questions, listed in the *Appendix*, concerning self-perceptions of general health, wellness, and quality of life. These questions were developed with reference to previously used measures, and in consultation with a psychometrician. They were designed to elicit self-reported changes in health, wellness, and quality of life using perceptions of scores "presently" against the explicit comparison of retrospectively recalled perceptions "before Network."

Serving as a point of comparison, 14 of the 55 items were taken from a psychometrically grounded "overall quality of life" instrument used by Woodruff and Conway (1992),[12] and adapted from the original landmark studies of Andrews and Withey (1976),[14] and Caplan et al. (1984).[15] In this instrument, to assess quality of life in a wide variety of areas, item responses were presented in a 7-point scale in Likert format[16] with choices scaled from 1-7 as "terrible," "unhappy," "mostly dissatisfied," "mixed," "mostly satisfied," "pleased," and "delighted."

A battery of 41 original items were partly adapted from a standard instrument assessing psychological status, "Structured Clinical Interview for DSM-III-R-Non-patient Edition."[17] Among these was one set of 10 items for evaluating stress (perceived ability to cope with the demands of the environment) relative to several spheres (e.g., family, work, etc.), using a 5-point Likert scale with choices ranging from 1-5 as "none," "slight," "moderate," "pronounced," and "extensive." An additional set of 11 original items assessed what was labeled "life enjoyment," wherein respondents were asked to rate their feelings and experiences contributing to a broad sense of inter- and intra-personal enjoyment on a degree scale of 1-5. Choices included "not at all," "slight," "moderate," "considerable," and "extensive." Finally, using a frequency scale of 1-5 representing "never," "rarely," "occasionally," "regularly," and "constantly," respondents were asked to rate their physical state (symptoms) and mental/emotional state (feelings, satisfaction) with a set of 10 items each.

Scale items were reverse coded, as necessary, to consistently reflect a higher score as indicative of better health. The items within each of the four domains were used to construct the indices assessing self-rated health.[18] In order to facilitate comparison across scales varying in number of items and response codes, each was re-scaled so that the theoretically lowest scores (i.e., those in which a respondent gave the lowest possible response to all items) were coded 0 and the theoretically highest scores were coded 1.[a] If a given individual answered at least half of the items within a scale, a missing (non-response) value was replaced with their mean response score for the questions they answered.[18] Finally, a *combined wellness scale* was constructed by summing the four wellness domains of physical state, mental/emotional state, stress evaluation, and life enjoyment, and then re-scaling in the 0-1 metric range. Respondents with missing values (n = 222) on any of these four scales were excluded from further analyses.[b]

Statistical analyses

The first concern, dealing with the development and initial validation of wellness-specific outcome measures, was to determine how well the item indicators (specific questions) and summated health domains indices, represented the theoretical concepts they purported to measure. That is, to what extent did the items/scales serve as valid measures of self-rated wellness and quality of life. Because reliable measurement is a necessary condition in building a case for validity, the first step involved analyses of reliability (essentially repeatability), or the extent to which the measures consistently yield the same results on repeated trials, free of "random error" or chance fluctuations. The more direct test-retest approach not only requires multiple measurements on the same individuals over time, and the often untenable assumption that the trait itself is stable, but is also entangled with any true differences in health associated with treatment (e.g. Network Care). Thus, the more commonly accepted analysis of internal-consistency reliability[19] was used as an indicator of how well the individual items (questions) within a scale reflected a *common* underlying health domains construct. This analysis is based on the logic that confidence can be increased, that the underlying theme of health/wellness is reliably measured by using multiple measuring instruments (questions), and gauging the extent to which the item scores are interrelated beyond random fluctuation. Chronbach's coefficient alpha is the most often used statistic,[19] and the most conservative test, because it is based on the assumption that all questions contribute equally to the measurement of the single wellness theme. This stringent requirement of "parallel measurement" means that alpha is a lower bound for the reliability of a multi-item scale.

A closer estimate of true reliability is coefficient theta, which relaxes the assumption of parallel measurement. This is considered a maximized alpha coefficient. Also, within the construct of internal-consistency reliability, theta is calculated from the results of principle components factor analyses, which are methods for discovering clusters of interrelated variables. The analysis essentially evaluates the extent to which the items share variability in concert with one another; this "shared variance" represents the underlying construct which, in this case, is the wellness theme of the measures.[20] The primary portion (principle component) of the shared variance among the items (eigenvalue) is a summary of how well the scores on given items account for, or predict, the scores on all other items (factor loadings). Calculation of the reliability coefficient theta is based on the magnitude of this principle component. Both alpha and theta coefficients are a product of the number of items and the strength of their intercorrelations, ranging from 0-1.00, with a cut-off of greater than 0.7 as the widely accepted rule of thumb for demonstrating internal consistency.[21]

In addition to its use in the reliability analysis, factor analysis is also a useful tool for assessing the validity of empirical measures.[20] The pattern of residual co-variation among the variables and factors, after accounting for the common variability in the principle component, was inspected to investigate how well the individual items support the theoretical theme (i.e., wellness domains). In addition, the convergence of the health domain scales and the psychometrically validated overall quality of life index was considered (i.e., reliability, inter-scale correlations, and associations with other available data such as spinal injury, etc).

Using these newly developed health domain scales, the third

objective of this study was the evaluation of retrospectively perceived changes in health and quality of life for patients while under Network Care. This was accomplished with bivariate comparisons of self-rated health scores "before Network" and "presently." The statistical significance of this difference was determined using two-tailed, paired sample t-tests, which tests the null hypothesis that any differences between individuals' paired "before Network" and "presently" scores result from chance fluctuation. Insofar as this evaluation involved the use of multiple scales, the usual alpha level of 0.05 was divided by the number of scales (5) to correct for multiple comparisons, yielding p=0.01 at a 99% confidence interval.[22] When the "p" value of this probability did not exceed the 0.01 cut-off, the null hypothesis was rejected, indicating that the difference is statistically significant. That is, at this level, only 1% of the time would the results be due to chance alone, or 99% of the time would the same pattern be found in the population from which the sample was drawn.[c]

In addition, difference scores were calculated within each of the four health domain indices between "presently," and "before Network," to determine a patient's perceived change. Summation of all health domain index scores represented a *combined wellness scale* for both "presently" and "before Network." The difference score of the *combined wellness scales,* herein referred to as the *"wellness coefficient,"* ranged from -1 to +1, with zero representing "no change," positive values indicating improved health, and negative values a worsening.

While the large size of this sample assures that a statistically significant effect will not be missed (due to sample error introduced by small sample size), statistical significance alone does not provide insight into the strength of a bivariate relationship or effect. Thus, measures of clinical significance, or size of "treatment" effect, are commonly reported.[24-25] In this regard, effect size was used to measure the magnitude of clinical or meaningful change in the present study. Effect size represents a standardized "benchmark" measuring the magnitude of clinical change.[26] Cohen (1977) provided the widely accepted definition of an effect size of 0.20 as small, 0.50 as moderate, and 0.80 or greater as large, with large being a change of magnitude at least four-fifths of a standard deviation of the baseline measure.[27] The most commonly accepted formula was used for calculating effect size: $(M_2 - M_1)/S_1$, where M_1 and M_2 are the group means at time one ("before Network") and two ("presently"), and S_1 is the standard deviation at time one.[24] This statistic compares the average differences in individuals' scores "presently" to the amount of deviation across the scores of all respondents "before" beginning care, with the idea that meaningful variations should be valued against the normal range of the majority of individuals' initial scores.

To better visualize the extent and meaning of clinical effects (effect size) in the sample, the percentages of patients whose perception of health/wellness worsened or improved more than 0.5 standard deviation, or remained stable, were calculated. The specific cutoff values for improvement thus vary as a function of the distribution and variability of each index, and generally reflects a conservative moderate effect size. These cut-offs were superimposed on histograms of the difference scores (perceived change) for each index.

Finally, to investigate if the interval between "before Network" and "presently" varied across respondents according to the duration of Network Care, the characteristics and associations of this parameter were examined. To capture meaningful time-bound differences in wellness, this continuous variable was categorically re-scaled to 1-3 months, 3-12 months, 12-36 months, and greater than 36 months. The categorization was based on the assumption that the influence of the length of time since beginning care does not operate in a linear (one-to-one) fashion with regard to either retrospective recall or the potential wellness related influence of care. Focusing on the validation of the retrospective recall approach, these four duration groups were compared in terms of the internal-consistency reliability of the self-rated health and quality of life indices, as well as other potentially relevant health characteristics. The connection between duration of care and the magnitude of retrospectively perceived changes in self-rated health and quality of life (score for "presently" minus the score for "before Network") was then examined using analysis of variance (ANOVA).[28] This statistic compares variation across the group means (defined by length of time under care) to the average variation within those groups; the resulting probability indicates the extent to which differences across groups are real or due to chance. In addition, measures of the clinically meaningful variation in wellness, i.e., effect size, are particularly well suited for comparing differences across these treatment duration groups.

Results

Patient Sociodemographic and Health Care Characteristics

The sociodemographic characteristics of the study population are presented in Table 1. The average age of respondents was 43 ± 12 years (mean ± SD) with a range of 18 to 95 years (U.S. mean age = 35 years). There was an over-representation of female (73%) compared to male (27%) respondents, and white ethnicity (94%), with the population reflecting only 1% black and less than 5% representation of all other ethnic groups. Socioeconomic characteristics of these respondents tended to be skewed toward higher education levels (79.6% college or graduate school), professional/white collar (69.8%) occupations, and higher income (Table 1).

A summary of the health histories for the population is given in Table 2. The majority of respondents self-reported their current physical and emotional health as good (64%-68%) or excellent (23%-28%), and only a few rated their physical/emotional health as poor (4%). Not surprisingly, in view of the wellness perspective of Network Care, the population had a low utilization of orthodox medical services, even though there was a relatively high incidence of persistent ailments (58%) and prior injury to the spine (47%). In this regard, only 28% of respondents reported seeing a physician for other than routine physicals; the last visit to a physician averaged 15 ± 30 months previous to completing the questionnaire. Fewer than half of the respondents (38-41%) reported taking any prescription or non-prescription medications currently, or for a duration of at least 2 months at any time in the past.

Patients had been under regular Network Care for an average of 21 ± 27 months (Table 2). The frequency of Network

Care was consistent across offices ranging from an average of 2.7 ± 1.3 times/ week during the first 2-3 months, to 2.4 ± 1.3 times/ week 3-6 months after initiating care, and further reducing to an average of 2.1 ± 1.2 times/ week after 6 months of care. Also consistent across offices, the vast majority of patients (75%) had been under previous chiropractic care for an average of eight years (96 ± 112 months) prior to beginning Network Care. In regard to Network Care, 95% of respondents reported that their expectations had been met, and 99% reported that they would continue care.

Table 1. Sociodemographic Characteristics of Patients undergoing Network Care

Number of respondents:		2,818
Age (years):		43.4 ± 11.5 (2770)
Gender:		(2799)
	Male	26.6%
	Female	73.4%
Ethnicity:		(2750)
	White	94.1%
	Black	1.2%
	Asian	1.2%
	Hispanic	1.7%
	American Indian	0.4%
	Other	1.4%
Marital status:		(2551)
	Single	31.1%
	Married	46.5%
	Divorced/Widowed	22.3%
Number of children:		(2615)
	None	46.3%
	One	15.6%
	Two	21.3%
	Three or more	16.8%
Occupation:		(2689)
	Blue collar	7.5%
	White collar	22.2%
	Professional/technical	47.6%
	Student	6.5%
	Homemaker	5.9%
	Retired	5.7%
	Unemployed	1.6%
	Self-employed	3.0%
Education :		(2785)
	High school	12.3%
	Other (avocational)	8.1%
	College	50.3%
	Graduate school	29.3%
Income: (median = $25-34,999)		(2542)
	< $24,999	40.0%
	$25-34,999	18.9%
	$35-44,999	13.5%
	$45-59,999	12.6%
	> $60,000	14.9%

Table 2. Self-reported health history of patients undergoing Network Care.

I. MEDICAL		
1. Current physical state:		(2794)
	excellent	25.0%
	good	71.0%
	poor	4.0%
2. Current emotional/mental state:		(2786)
	excellent	29.0%
	good	66.0%
	poor	5.0%
3. Persistent ailments:		(2684) 58.0%
4. Physician visits (other than routine):		(2787) 28.0%
5. Last medical visit (months):		(2324) 15.0 ± 29.6
6. Currently taking medications:		(2214) 41.0%
7. Medications taken for at least 2 months in past:		(1916) 38.0%
II. CHIROPRACTIC		
1. Spinal injury:		(2737) 48.0%
2. Time since spinal injury (yr.):		(1075) 15.5 ± 17.4
3. Duration of Network Care (mo.):		(2510) 21.4 ± 27.0
4. Prior chiropractic care:		(2774) 75.0%
5. Duration of prior care (mo.):		(1506) 93.2 ± 111.6
6. Frequency of Network Care:		
(appointments/wk)	initial 2-3 mo. (2240)	2.7 ± 1.2
	initial 3-6 mo. (1812)	2.5 ± 1.4
	> 6 mo. (1539)	2.2 ± 1.1
7. Expectations met with Network Care:		(2367) 95.0%
8. Number choosing to continue Network Care:		(2770) 99.0%

Self-Rated Health, Wellness and Quality of Life Scale Item Analyses

In this study, the efficacy of Network Care was assessed in terms of several newly developed self-rated health indices, as well as the overall quality of life index, which were rated by the respondents both "presently" and retrospectively recalled "before Network." The means and standard deviations for the four domains of health, overall quality of life, and combined wellness scales (indices) "presently," and "before Network" are presented in Table 3. The re-scaled means of the four self-rated health domain scales "presently," ranged from 0.67-0.70, corresponding to a rating of about 70% of the maximum possible in the 0-1 metric. Mean scores "before Network" hovered around the mid-

Table 3. Reliability Coefficients and Means for Self-Rated Health, Wellness and Quality of Life Scales (N ≥ 2596)

INDEX	"Presently"				"Before Network"			
			Internal Consistency Chronbach's				Internal Consistency Chronbach's	
	Mean	St Dev.	Alpha	Theta	Mean	St Dev.	Alpha	Theta
Combined Wellness	.678	.100	.8949	.9110	.507	.138	.9230	.9342
A. Physical State	.701	.121	.7418	.7539	.558	.154	.7581	.7670
B. Mental/Emotional State	.666	.136	.8210	.8310	.494	.190	.8612	.8680
C. Stress Evaluation	.674	.155	.8176	.8336	.483	.199	.8392	.8528
D. Life Enjoyment	.669	.135	.8367	.8490	.502	.145	.8309	.8360
Overall Quality of Life	.700	.142	.9297	.9396	.542	.171	.9447	.9504

point of 0.5, or about 50% of maximum in the same metric, corresponding to perceptions of "moderate/occasionally" in the original response scale. Across the separate scales, respondent variability (standard deviation) was approximately 10% greater "before Network" than "presently," as would be expected given the cognitive complexity of such recall. The mean for the overall quality of life index is higher on average (with similar variability), but not dramatically different than those of the separate health domains indices for "presently" or "before Network," or the *combined wellness scale* (summation of all health indices).

The results of the scale item reliability analyses are also presented in Table 3. For all indices, both "before Network" and "presently," Chronbach's alpha coefficients are clearly above acceptable levels (i.e., > 0.7), indicating strong internal consistency, that is, an interpretable underlying theme for the scale items of each wellness domain and overall quality of life. Only for the stress evaluation scale did further inspection result in the deletion of an item (stress associated with school) to optimize reliability. For the remaining 54 items, the deletion of any variable resulted in lower reliability for the respective scale, indicating that each provided uniquely important information. Moreover, given that the alpha for the *combined wellness scale* was substantially higher than the coefficients for the separate domains, it is apparent that this overall collection of questions meaningfully reflects a single theme labeled "wellness." The alpha coefficient for the *combined wellness scale* (as well as those for the separate health domains) compares favorably to the expected higher alpha for the established overall quality of life index.

To further explore the reliability and validity of these scales, principle components factor analyses of the respective items was conducted. Based on resulting calculations (derived from the first eigenvalue [Methods]), in all cases coefficient theta indicated that the items in all scales were roughly parallel (equal) and that factor weighted scaling was not necessary as it would not produce substantially more reliable scales. The set of items in each scale met the established criteria supporting the measurement of a single underlying phenomenon. The first component theme accounts for a large proportion of the variability in the items (as indicated by the magnitude of theta) with gradual decreases across subsequent components. In addition, all items had factor loadings (item contributions to the overall theme measured by the scale) of greater than 0.4 (exceeding a 0.3 cutoff) on the first component. Moreover, factor loadings of the items indicating their contribution to subsequent components, were less than the contributions to the primary (principal) factor.[19]

To address the relationships between the scales, inter-scale Pearson correlation coefficients both "before Network" and "presently" were examined. Coefficients for the separate wellness domains of physical state, mental/emotional state, stress evaluation, and life enjoyment revealed moderate to substantial correlations, ranging from 0.15 (between physical and enjoyment "before Network") to 0.67 (between stress and emotional state "presently"). All scales showed slightly weaker inter-scale correlation in the "presently" than the "before Network" scores. Overall, the magnitude of these correlations suggests that while the separate domains are meaningfully related, they were not redundant items/scales. The correlations between the overall quality of life scale and the four health domain scales ranged from 0.34 (with physical state) to 0.58 (with life enjoyment) "presently," with a similar pattern of higher values "before Network" (0.42 - 0.68). The correlations between the overall quality of life and combined wellness scale were 0.74 "before Network" and 0.66 "presently," again suggesting that they are meaningfully related but not redundant.

To further explore the characteristics of this *combined wellness scale,* both "before Network" and "presently," factor analysis combining all items from the four health domains was conducted. The results support the conclusion that the combined set of all items measured a single phenomenon or "wellness theme" with factor loadings consistently exceeding 0.3 (criteria described above); this conclusion also derives from the fact that the theta reliability of the *combined wellness scale* is higher than the reliability of the separate health domain scales. However, further inspection of the factor scores did reveal a pattern wherein the items within the respective domains tended to cluster together identifiably after removing the primary variance. Thus, even though the *combined wellness scale* is slightly more reliable and more strongly in accord with the overall quality of life index, additional information is available by considering the (sub-) scales separately. In addition, there is some statistical evidence, in the magnitude of the second factor scores, that the life enjoyment scale may represent a second theme somewhat distinct

Table 4. Perceived Changes (Difference Scores) in Self-Rated Health/Wellness and Quality of Life Scales Between "Before Network" and "Presently" (N ≥ 2596)

INDEX	Difference Scores Mean	St Dev	Improved*	Paired t-test p value	Effect Size
Combined Wellness	+.171	.136	76.4%	.000	1.24
A. Physical State	+.144	.128	63.9%	.000	0.94
B. Mental/Emotional State	+.173	.163	64.3%	.000	0.91
C. Stress Evaluation	+.194	.196	66.2%	.000	0.98
D. Life Enjoyment	+.167	.164	68.9%	.000	1.15
Overall Quality of Life	+.158	.168	58.9%	.000	0.93

* Positive change greater than ± 0.5 standard deviation, corresponding to a moderate or greater effect size.

from that shared by the physical state, mental/emotional state, and stress evaluation constructs. Accordingly, results are presented for both the separate health domains and the *combined wellness scale* in the remainder of this report.

Retrospective Outcomes Assessment.

These scales were used for the evaluation of retrospectively perceived changes in self-rated health and quality of life for patients under Network Care (Table 4). For each individual, perceived change was calculated as the difference between "presently" and "before Network" scores. For every outcome measure, the mean difference score was positive, indicating that for all scales, on average, there was reported improvement with care. The perceived change (difference) in scores of the *combined wellness scale ("wellness coefficient")* indicated an average of + 0.17 in a range of -1 to +1. This was interpreted as an overall increase in wellness as the value was consistent with statistically significant differences in the scale scores and positively (+) signed (numerically closer to +1 than -1). In regard to the separate health domains, the differences between "presently" and "before Network" were also all positive and in the same range as the *wellness coefficient* (Table 4). The magnitude of these perceived changes varied from +0.14 to +0.19, which compares to a +0.16 increase for the overall quality of life index.

The results of two-tailed paired sample t-tests indicate that these perceived changes (i.e., differences between "presently" and "before Network" scores) were statistically significant for all outcomes ($p<0.000$); demonstrating that the differences in scores were not due to chance fluctuations. The effect size statistics further showed that these improvements (ranging from 0.91 to 1.15 across the separate scales) were above 0.8, indicating that a large positive clinical outcome had occurred across all domains. In addition, there was notable correspondence between the separate health domains derived for the purpose of this study and the standardized overall quality of life index (Table 4).

The advantage of retaining the separate domains of physical state, mental/emotional state, stress evaluation, and life enjoyment, as well as the *wellness coefficient*, becomes apparent in comparing range of effect sizes for these outcomes. Retrospectively perceived changes associated with physical state, mental/emotional state, and stress evaluation are comparable in magnitude with overall quality of life assessments (range=0.9-1.0). In contrast, the change in life enjoyment is notably greater (1.15). Moreover, when all scales are combined in the *combined wellness scale*, the effect size calculation shows a large (1.24) positive clinical effect associated with Network Care (Table 4).

To better visualize clinical improvement in outcomes, Figure 1 illustrates the self-perceived change histograms for each index, with each data point in the histogram representing the difference score for an individual. The cut-offs for the categorization of each respondent as worsened, no change, or improved are shown superimposed, and ranged from ± 0.07 for life enjoyment to ± 0.10 for mental/emotional state and stress evaluation. Across all scales, 4% or less of respondents reported decreased levels of self-rated health, wellness, and quality of life, and between one-fourth and one-third did not change beyond one-half standard deviation within the scale. In contrast, as also noted in Table 4, the majority of respondents reported at least a moderate clinical improvement across all outcome assessments. In spite of the conservative cut-off values, 59% improved on the overall quality of life index, and about two-thirds of respondents showed clinical improvements in physical state, mental/emotional state, stress evaluation, and life enjoyment indices, with over 76% perceiving improvement in the wellness coefficient.

Analyses of Outcomes by Duration of Care

While the variable representing length of time since beginning Network Care in months, i.e., duration of care, was significantly skewed toward the lower range, the re-scaled version representing duration intervals of 1-3, 3-12, 12-36, and 36+ months was normally distributed (i.e., not significantly skewed). Table 5 presents the means, standard deviations, and effect sizes "presently" and "before" Network Care for each of the four health scales, the *combined wellness scale*, and overall quality of life, for the four duration of care intervals.

ANOVA results showed that the mean "presently" scores differed statistically ($p<0.003$) across the duration groups for all self-rated health and quality of life scales, as well as the *combined wellness scale*, indicating that increasing duration of Network Care is significantly associated with increasing levels of perceived health and wellness. Although this pattern is meaningful in itself, the retrospectively recalled "before Network" scores in the four health domains, the *combined wellness scale*, and overall quality of life also provided the opportunity to examine the relationship between possible perceived changes in health and wellness and length of time since starting care; i.e., how stable is the

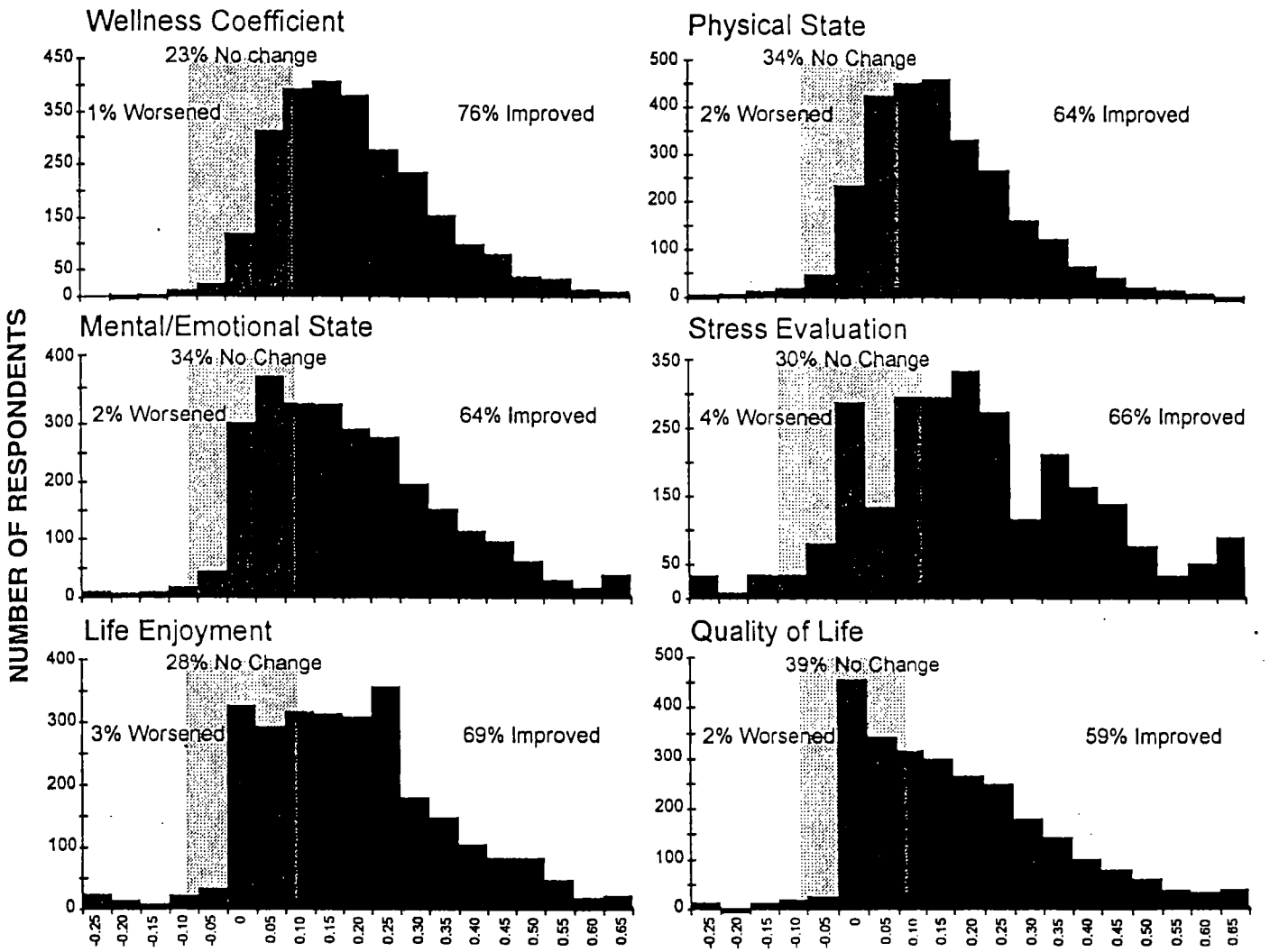

Figure 1 illustrates the self-perceived change histograms for the wellness coefficient, each health scale, and overall quality of life. Data points in the histogram represent the difference scores for each individual (e.g. difference between "presently" and "before Network") for each scale. The percentages of patients who worsened or improved more than one half of the standard deviation (± 0.5 SD.), or had no change (stippled areas in figure) are superimposed over each histogram. The specific cutoff values for improvement thus vary as a function of the distribution and variability of each index. The cut off values ranged from ± 0.06 for physical state to ± 0.10 stress evaluation.

reliability of retrospective recall of one's health given varying time intervals?

To begin to address the validity of the retrospective recall strategy, the within group variation (standard deviation) in "before Network" scores in all scales was examined (Table 5). This differed only minimally across these duration groups, and showed no pattern of increased variability with increased duration of care (i.e., length of recall). Moreover, the reliability coefficients were strikingly similar across duration groups, varying in magnitude by 1% or less for all scales (except physical state at 2.5%), and showed no trend toward lower reliability associated with increased length of retrospective recall. Thus, a longer interval since "before Network" is not associated with increased random error in the measurement of quality of life, self-rated health, or wellness in these data.

ANOVA results indicated that the mean "before Network" scores for overall quality of life, stress evaluation, life enjoyment, and the combined wellness rating were significantly different across the duration groups ($p < 0.003$). In each instance, the "before Network" wellness scores were lower with longer intervals in care.

The *wellness coefficient*, being the difference between "presently" and "before Network," reflected in one measure the incrementally lower "before Network" wellness ratings concomitant with the progressively higher "presently" scores, observed with longer intervals in care. Moreover, the self-rated health scales, *combined wellness scale*, and overall quality of life across the duration groups showed significant positive differences (improvement). As is readily apparent in Figure 2, respondents within groups, defined by increasingly longer duration of care, report-

Table 5. Duration of Network Care Reliability Analyses for "Before Network" and "Presently" Self-Rated Wellness and Quality of Life Scales (N ≥ 2330).

	"Presently"		"Before Network"			Effect Size
	Mean	St Dev	Mean	St Dev	Alpha	
Combined Wellness						
In Care 1-3 mo.	.665	.102	.526	.136	.953	1.03
3-12 mo.	.669	.096	.517	.135	.951	1.13
12-36 mo.	.680	.099	.495	.139	.956	1.33
> 36 mo.	.702	.098	.482	.139	.954	1.54
A. Physical State						
In Care 1-3 mo.	.691	.123	.568	.152	.773	.81
3-12 mo.	.695	.115	.562	.153	.747	.87
12-36 mo.	.705	.119	.551	.156	.841	1.00
> 36 mo.	.717	.124	.555	.154	.754	1.08
B. Mental/Emotional State						
In Care 1-3 mo.	.650	.144	.508	.187	.859	.76
3-12 mo.	.653	.135	.497	.185	.856	.84
12-36 mo.	.673	.134	.489	.191	.867	.97
> 36 mo.	.686	.130	.476	.196	.858	1.09
C. Stress Evaluation						
In Care 1-3 mo.	.670	.153	.514	.196	.834	.79
3-12 mo.	.668	.152	.497	.196	.838	.89
12-36 mo.	.667	.156	.462	.199	.841	1.10
> 36 mo.	.698	.150	.462	.199	.838	1.24
D. Life enjoyment						
In Care 1-3 mo.	.646	.135	.517	.150	.847	.85
3-12 mo.	.656	.138	.516	.143	.836	1.00
12-36 mo.	.677	.135	.489	.143	.838	1.13
> 36 mo.	.700	.128	.478	.139	.816	1.59
Overall Quality of life						
In Care 1-3 mo.	.670	.147	.561	.177	.942	.61
3-12 mo.	.693	.136	.566	.165	.942	.78
12-36 mo.	.704	.145	.524	.174	.948	1.00
> 36 mo.	.735	.134	.518	.168	.943	1.31

ed incrementally higher positive *wellness coefficients*. That is, not only were the changes between "presently" and "before Network" scores statistically significant for all outcomes, but the extent of that change in the *wellness coefficient* is positively associated with the duration interval of care. Moreover, the percent of respondents exhibiting a greater than moderate clinical effect also increased with increased duration of care.

To further explore the association between wellness and duration of care, two-tailed paired sample t-tests were conducted on the change between "presently" and "before Network" scores for the *combined wellness scale*, and overall quality of life, separately within each duration of care group. Regardless of the time interval since beginning Network Care, the mean combined wellness improvement was statistically significant. Additionally, progressively higher effect sizes were evident across the four categories of duration of care in nearly proportional increments; 1.03 for 1-3 months, 1.13 for 3-12 months, 1.33 for 12-36 months, and 1.54 for 36+ months of Network Care. A similar pattern was obtained for the overall quality of life index (Table 5).

These proportionally increasing effect sizes (about 0.1 per year of care) suggest that these particular intervals are meaningful time categories, perhaps connected to the accrued benefits of care. Thus, even for respondents who began care within three months of completing the survey, the statistic is well above the 0.8 value indicating a large clinical treatment effect; for those in care more than 3 years, the effect size is nearly double this benchmark. This systematic difference reflects clinically meaningful retrospectively perceived improvements in self-rated

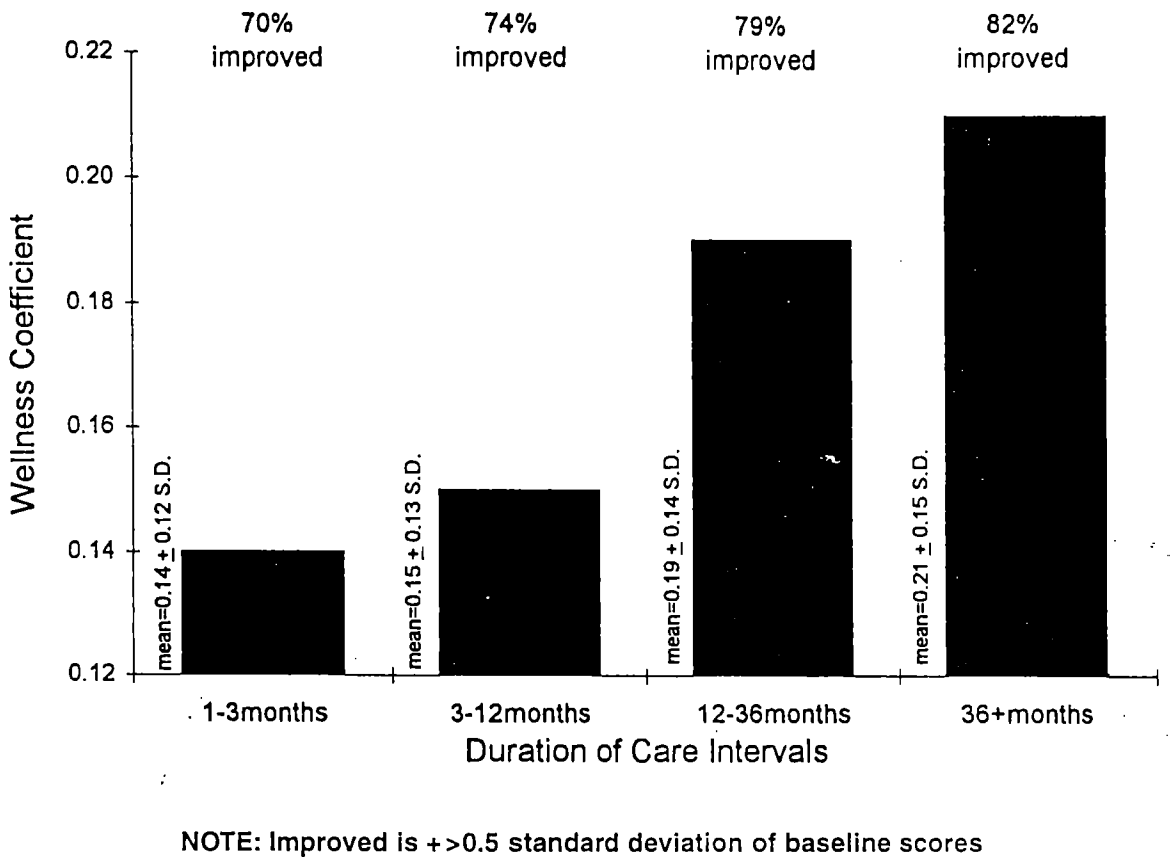

Figure 2 shows mean wellness-coefficients versus duration of care. Respondents within each group, defined by increasingly longer duration of care, reported incrementally higher, positive wellness coefficients. The percent of "improved" respondents, defined as those greater than one half of the standard deviation of each duration of care category, are also indicated; note that the percent "improved" also increased in each of the duration of care intervals. These percentages represented a greater than moderate clinical effect.

health, wellness, and overall quality of life, associated with Network Care.

Discussion

A primary objective of this study was the evaluation of retrospectively perceived changes in self-rated health and quality of life for patients undergoing Network Care. The results from analyses of 2,818 respondents has provided compelling evidence for the beneficial effects of Network Care. This population of patients undergoing Network Care self-rated their status "presently" at a level of 0.7 (70%) in a metric of 0 - 1.0, and were "mostly satisfied" with their quality of life. This was a statistically significant increase over their self-reported wellness "before Network" which averaged about 0.5 (50%), and, on average, had "mixed" satisfaction about their overall quality of life. Over 76% of respondents reported positive, clinically significant *wellness coefficients*, incorporating perceived changes in physical state, mental/emotional state, stress evaluation, and life enjoyment. Categorizing patients in terms of duration of care revealed progressively higher reported effects with time under Network Care.

The precise mechanisms underlying these effects are uncertain, but it is hypothesized to be, in part, associated with changes in levels of circulating factors released by the pituitary-adrenal cortical axis. While this hypothesis will be tested in subsequent studies, by analogy, such changes have been associated with the relaxation response during meditation[2,3,6] and stress-reduction programs.[4,5,7]

The Survey Population

This first census of active practitioners of Network Care realized a 49% participation by ANC practitioners. Given the practitioner participation rate, coupled to the strong statistical power of the findings, conditions contributing to a positive bias could be operational. However, respondents to a separate practitioner survey (conducted during the same time frame) indicated that among 116 respondents, 97% followed the suggested protocols outlined for Network Care. This implies that Network Care is practiced consistently. Moreover, since the patient responses were received from 156 different practices, widely distributed across the range of practice locales, there is no prevalent rationale which would suggest that practitioner participation was biased. Nevertheless, further study investigating the impact of

practice style on patient outcomes will be necessary to better understand this issue. On a practical note, if indeed the sample is biased towards results obtained by practitioners more committed to health care delivery, then the present results should be viewed as a "benchmark" against which new practitioners could set practice goals and objectives.

Based on population parameters and inclusion criteria, the estimated survey response rate of 67-71% is generally considered to be "very good" particularly for mail surveys.[39] Although the overall response rate is in the acceptable range, it is important to ascertain as much information as possible regarding non-respondents in order to detect other systematic bias. The telephone follow-up with practitioners confirmed some reluctance to participate on the part of patients for reasons of time constraint, privacy and/or disinterest in the study, but not for reasons that might bias the results (e.g., only the best patients, or those receiving the greatest amount of care, etc.). It is difficult to estimate the number of patients declining to participate for reasons of adverse reactions or poor outcomes. However, among those responding, there was a small number (n = 29, 1%) who indicated that they would not be continuing Network Care. Those choosing not to continue care reported the smallest (but still significant) improvements in the indices for physical state and mental/emotional state, although no significant change from base line was detected in this group with regard to the remaining three indices (stress evaluation, life enjoyment, and quality of life). While it is believed that "self selection" bias is not a significant feature of the present study due to its wide range of participating practices, a longitudinal study is currently underway to assess patients from the onset of care through a specified time period. This approach will provide the opportunity to ascertain if only those receiving positive benefits from care "self select" as survey respondents by evaluating non-respondents, as well, relative to their clinical benefit status.

The first objective of this study was to characterize the population under Network Care. The vast majority of the population studied are from offices within the United States (93.5%), with a small representation from Canada (4.6%), Australia (1.0%), and Puerto Rico (0.8%). Given the international scope and other inclusion criteria defining this population, it is difficult to relate directly to the U.S. population. Although the age of respondents appears to be normally distributed, the average of 43 years is higher than the national mean (35 years),[40] and is likely because of the exclusion of patients less than 17 years of age as well as the lower incidence of health care utilization in early adulthood. There were clearly more female than male respondents, and the population was predominantly white. Socioeconomic characteristics of these respondents tended to be skewed toward higher education levels, professional/technical occupations, and higher income. These patterns have also been noted in studies of alternative/complementary medicine in the U.S. and Europe,[41-43] as well as studies of chiropractic care.[44-45] Whether or not the results of this study extrapolate to the general population awaits further investigation.

A more interesting challenge is to attempt to understand the factors contributing to the uniqueness of the population under Network Care. The cost of care may be one factor. There is a high incidence of self-pay with patients undergoing Network Care. The fee for an office visit varies greatly depending on practice size, location, pay plan (individual, family, monthly, weekly, etc.), but ranges from approximately $15-$50 per visit. Given the frequency of care, i.e., ca. 2 times/week (Table 2), this amounts to about $120 to $400 per month. While this is consistent with the approximate amount paid by Americans for health insurance and out-of-pocket expenses for alternative/complementary care,[41] other studies document the strong negative correlation between family income and access to health care,[46] and by analogy, cost may restrict access to Network Care. In terms of ethnicity, the lower family incomes among ethnic minorities may account for the limited access of these groups to the health care structure in general.[47] Other factors such as cultural attitudes and health belief structures are strong determinants of access to health care,[48-49] and need to be examined within the context of Network Care. Finally, the high incidence of females in the present population (73%) relative to the general population (51%) is consistent with health care utilization patterns in general[50-52] and is currently undergoing further study.

Study Design

Although there are a number of objective criteria for assessment of vertebral subluxation such as surface electromyographic recordings and thermography, which may be applied to the analysis of patients under Network Care, these objective criteria do not take into consideration the overall health and wellness status of the patient. For this reason, it was important to assess the broader issues by use of self-rated health, and overall quality of life (see Introduction for details).

Self-rated health measurements are used with increasing frequency as measures of primary and secondary outcomes in clinical studies, and in a growing number of studies, have been shown to be an impressive, independent predictor of outcomes, including mortality and longevity.[10] All self-reported survey data is subject to potential response effects, such as when respondents provide socially desirable responses, or are inclined to respond similarly to all items (common method variance). In the present data, these issues do not appear to be particularly problematic: arguing against socially acceptable responses is the observation that actual responses included the range of possible scores, and the issue of common method variance error is negated because summated scales showed greater variability than the separate original items. Also, analyses of cases with missing values on the outcome assessments did not reveal non-response to be a systematic biasing factor.

In the present study, self-rated health was used as an evaluative instrument to measure perceived change in patient perceptions, "presently" relative to "before Network." The patients' self-reported assessments of their prior status before initiating Network Care allowed analysis of perceived trends over time, but not specific cause-effect relationships. Although the cross-sectional design holds advantages for the purposes of this study, longitudinally collected data is necessary to verify time-related changes (e.g., with duration of care).

The ability of an evaluative instrument to detect change (i.e., responsiveness) can be compromised by floor/ceiling effects in which patients with the best scores may continue to improve

(beyond the scope of the response range), or the health status of patients on the low end of the scale continue to worsen. No floor effects were encountered. Although examination of the data revealed potential operation of ceiling effects, further investigation showed less than 2% of cases could be considered suspect (>95% of maximum), with trivial attenuation of outcomes.[d] Thus, this wellness survey demonstrates sensitivity for measuring improvements in an already healthy population.

Validation of survey instrument

The second objective of the study was the development and validation of a wellness-specific survey instrument. Reliability analyses showed high levels of internal consistency for the physical state, mental/emotional state, stress evaluation, and life enjoyment scale items, indicating that each scale represents a dependable and interpretable measure of its respective theme. The notably, but not problematically, lower alpha coefficient for the physical state scale is not unanticipated given the relatively greater substantive diversity of questions in this scale; it is not necessarily expected that the presence of allergy, for example, will co-occur to a great degree with flexibility of the spine. Moreover, the high reliability of the *combined wellness scale* suggests that efforts to formulate a broad and coherent wellness assessment reflecting the WHO definition of health were realized. The reliability coefficients obtained in this sample for the overall quality of life scale were comparable to values obtained in prior studies with this instrument.[12,52]

Validation of retrospective recall

The strategy developed in this study of asking respondents to retrospectively recall their level of health and wellness before initiating Network Care as well as "presently," represents an integration of the classic pre-post study design with the increasingly recognized importance of assessing health interventions in terms of "patient-centered" outcomes. The differences between the standard deviations and reliability coefficients for the "before Network" scales compared to those for the same scales measured "presently," do not appear to be substantial, suggesting only the increased complexity of the cognitive process involved in recalling prior levels against the standard of present experiences. Moreover, the scale properties across the duration of care groups showed no trend toward greater variability or lower reliability associated with increased length of retrospective recall.

Retrospective Outcomes Assessment

The third objective of the present study involved a retrospective outcomes assessment using the data derived from the survey. Results show that statistically and clinically significant changes occurred within the respondent population regarding self-rated health outcomes and quality of life "before Network" and "presently." In this regard, the 70-76% of respondents reporting moderate to large improvement in the four scales of health assessment, and 59% of respondents reporting such improvement in overall quality of life, and effect sizes about 0.9 (clearly exceeding the benchmark for large clinical significance for every measure), substantiate the health promoting premise of Network Care.

Exploration of these perceived health improvements over given time intervals yielded suggestive evidence for the long-term benefits of Network Care. Across the four duration of care groups, progressively higher percentages of respondents (70%, 74%, 79%, 82%) reported clinically meaningful improvement in every measure. Moreover, clinical effect sizes across the four durations of care, averaging a range of 0.8 to 1.2 for the four domains of health assessment, and 0.6 to 1.3 for overall quality of life, attest to perceptions of consistently greater benefits of care as a function of duration of care. This finding also has implications in regard to the concept of maximum clinical benefit, which presupposes a "leveling off" effect.[11] On the contrary, the current findings regarding Network Care suggest that clinical benefits accrue over time, with no indication of a maximum in excess of three years of care.

Similarly, the *"wellness coefficient,"* representative of the difference between the *combined wellness scale* ratings "before Network" and "presently," increased systematically as a function of duration of care. This indicated a continuum of improvement in overall "wellness" while under Network Care, initiated even among those respondents who began care within three months of completing the survey. Even more striking were the proportionally increasing clinical effect sizes across the four duration of care intervals (about 0.1 per year of care), indicating improvement in every index of health measured in this study.

This continuum of improvement in the *"wellness coefficient"* not only reflected the progressively higher "presently" ratings across the four duration of care intervals, but also a progressively lower self-rating of "before Network" scores. This might suggest that those respondents who remained in care longer were in poorer health before initiating Network Care. However, other available information does not support this conclusion. In particular, the overall ranking of health (Table 2, Item I.1) and whether or not the respondent had ever injured their spine, or experienced a physical or emotional trauma, was not significantly associated with duration of care group.

What then accounts for the downgrading of self-reported health status as a function of the duration of retrospective recall? It appears that these broad, patient-centered health measures detect advancements in wellness such that respondents have a new standard against which to gauge their recalled "poorer" levels of health. The interesting exceptions to this pattern were the "before" scores for physical and mental/emotional state which were not significantly different across the four duration of care intervals. Perhaps these more concrete domains are less susceptible to perceptual shifts, and/or are perceived as less relevant to the overall experience of wellness. Further research in regard to the broad concept of health, expressed in the WHO definition, will need to examine the content validity, or differences and relationships between the various domains of health contributing to experienced wellness.

Construct Validity

In regard to self-reported health, as with other outcome measures such as pain, quality of life, and depression, there is no "gold standard,"[5] or universe of content accepted as totally adequate to define the quality being measured.[14] Thus, the process of instrument validation "requires a pattern of consistent find-

ings involving a number of different researchers using different theoretical structures across a number of different studies."[19]

While ongoing longitudinal studies will assess the present instrument on a more global basis, the process of estimating its construct validity has begun through the development of a conceptual basis for measuring the theme of "wellness." This was accomplished by incorporating aspects of the WHO definition of health including physical, mental, and social well-being. As anticipated, individuals provided a broad base of information by retrospectively assessing their health "before Network" relative to their present status under Network Care. This permitted conclusions to be drawn regarding perceived change in health through a *"wellness coefficient,"* representing the retrospectively perceived difference in the combined domains of all four reported health scales, which were shown to representatively support and validate a principal theme. Nevertheless, the results from factor analysis suggest that the life enjoyment domain (representing the most understudied aspect of the WHO definition) is somewhat unique in these data, and the effect size for this scale indicates that life enjoyment may be a particularly important aspect of health contributing to perceived wellness. Overall, the *wellness coefficient* as an outcome of Network Care demonstrates clear internal validity.

Since the instrument was comprised of scales surveying established indicators of health status, as well as incorporating a reference scale assessing quality of life, a basis for further evaluation of its external reliability and construct validity was provided. The *combined wellness scale* demonstrated convergent validity with the overall quality of life scale. That is, these separate and combined self-rated health scales produced results comparable to the overall quality of life scale, while the reported reliability coefficients are consistent with those for this same scale applied to a wide variety of subjects in other studies.[12, 14, 15, 52]

While validation of the survey instrument will require its continued application in a variety of settings, initial findings indicate the instrument has a high level of sensitivity in measuring the central theme of "wellness," is reliable, and exhibits internal and external construct validity.

Summary and Conclusions

Findings of the present study elaborate the importance of characterizing and investigating the efficacy of this non-medical health care practice in accord with its specific objectives. In that regard, evidence is provided which shows that Network Care:

1. Is utilized by a unique population, with socioeconomic, gender distribution, and educational characteristics similar to those seeking other forms of health care different from orthodox medicine.
2. Is associated with significant "retrospectively recalled" improvement in self-rated perceptions of health, wellness and overall quality of life.
3. Results in a large (>0.9) positive clinical effect in every health-related domain investigated.
4. Is associated with significant improvement in self-rated perceptions of "wellness," positively correlated with length of time under Network Care.

Within the boundaries of the study design, these findings provide substantial evidence that Network Care should be included among those practices with established health benefits.

Results of this study have been presented in a manner which permit comparison with data to be derived from continuing longitudinal study of Network Care, as well as future studies of other non-medical health related approaches. Moreover, the newly developed survey instrument, which has initially shown a high level of validity in measuring an underlying "wellness" theme, is an important contribution to further study of the holistic definition originally proposed by the World Health Organization. Repeated use of the survey instrument, or its component scales, by both the non-medical and medical communities, will serve to test its validity as a means through which information can be acquired linking self-rated perceptions of wellness to a variety of health issues.

These initial findings show that Network Care is associated with significant improvement in all indicators of health evaluated, and demonstrate a strong association between Network Care and self-reported, positive change in overall health/wellness. Although these findings are supportive of Network Care, they must be interpreted within the boundaries of a cross-sectional study, i.e., lack of test-retest responses and no control population. For this reason, longitudinal studies, with repeated measures should be undertaken to provide a more thorough assessment of changes over time. The next phase of research underway, involves a longitudinal study to provide such long-term assessment of patients and controls to evaluate dynamic behavior and perceptions of patients under Network Care. This longitudinal study will be combined with laboratory research involving EMG analysis, changes in stress-related hormones, immuno-chemical profiles, computerized platform posturography, and mathematical modeling to further elucidate the neurological/physiological mechanisms underlying Network Care.

Endnotes

[a] The re-scaling involved simple linear transformations. Each index was created by first summing the equally weighted items (ranging from 9-14) scores to yield composites with theoretical ranges between n (number of items) and n times p (the number of points on the Likert scale — 5 or 7). The transformation was of the form $I_{trans} = (I - n)/(pn - n)$, where I is the index in its original metric and I_{trans} is the transformed index.

[b] Using this conservative strategy, approximately 8% of cases were considered missing on this all domains wellness scale, with most missing on more than one sub-scale. While this is not a problematically high percentage, nevertheless, non-responders were compared to responders in terms of all otherwise available information. No significant differences were found in sociodemographic, health, or health/chiropractic care characteristics which would systematically bias the results of the outcomes assessment.

[c] In addition, because of a perceptible degree of positive skewness in the items/scales, which would violate the statistical assumption of normality, non-parametric versions of all bivariate statistics were also obtained. In every case, the numerical results were highly similar, and substantive results identical.

Based on a desired probability of 0.95 to find a 20% before 'presently difference (with a standard error of 2%) to be statistically significant at the 99% confidence interval, calculations indicate a minimum of 21 cases were required.[7] Clearly the bias of small sample size is not an issue in this investigation.

[d] The individual items making up the summated scales show some tendency toward positive skewness on the response code metric (mean scores above the

midpoint of the range), opening the possibility for the operation of these ceiling effects, which can attenuate the difference scores and thus the empirical magnitude of Network Care effects. One advantage of scaling the set of items is that the procedure can diminish this deviation from normality by capturing additional variablity across the set within each individual's scores. Among the "before Network" summated scales, only physical state and quality of life show this above midpoint mean; although all of the "presently" scales are more positively skewed toward wellness, the mean is nevertheless well below maximum. Moreover, the number of potentially affected cases (i.e., those with summated scale scores above 80% of maximum "before Network") was estimated to be at or less than 5.7% of respondents for the individual scales, and only 1.4% (36) cases for the all domains wellness scale. While the mean wellness change score was clearly lower for these small groups than for the sample overall (indicating the operation of ceiling effects), the number of these cases at the maximum presently score (0.95+) was less than 1% , with less than 2% of the overall sample scoring at or above this upper level on present wellness.

Acknowledgments

The authors wish to acknowledge and thank Dr. Ralph Boone for his tireless effort in the design and implementation of this project and his invaluable contributions to this manuscript. The authors are grateful to Dr. Susan Schmidt and Dr. Steven Stumpf for their assistance in the design and critical review of the wellness surveys. We gratefully acknowledge Dr. Donald Epstein, Dr. Jackie Knowles, Dr. Michael Stern, Dr. Jan Kirschner, Dr. Linda Capra, Dr. David Tribble, Dr. Richard Kaye, Dr. Laska Hospers and the ANC practitioners and their practice members for their untiring support in conducting the survey. We are also indebted to Elenea Post-Gallegos, Thomas Pak, David Vaught, Haitham Al-Sayed, Jeffrey Husain and Andrea Craig for their assistance with data entry and analysis. The authors also wish to thank Professor Howard Tucker, Thomas Kurosaki and Tom Taylor for their advice on statistical analysis. This project was funded by a grant to RHIB from the Association for Network Chiropractic.

References

1. Epstein D. Network Spinal Analysis: A system of health care delivery within the subluxation -based chiropractic model. Journal of Vertebral Subluxation Research Aug. 1996; 1(1)51-59.
2. World Health Organization: The first ten years of the World Health Organization. Geneva: WHO, 1958.
3. Macbeth HM, 1st ed. Health Outcomes: biological, social, and economic perspectives. New York, NY: Oxford University Press, Inc., 1996.
4. Wilson IB, Cleary PD. Linking clinical variables with health-related quality of life. A conceptual model of patient outcomes. JAMA 1995; 273 (1): 59-65.
5. McDowell J, Newell C. Measuring Health: A guide to rating scales and questionnaires. 2nd ed. New York: Oxford University Press, 1996.
6. Seventh Report to the President and Congress on the Status of Health Personnel in the United States. Washington: U.S. Department of Health and Human Services. Public Health Service Health Resources and Services Administration. Bureau of Health Professions, March 1990: XV1 - XV14.
7. Palmer BJ. The subluxation-specific - the adjustment-specific. Davenport, IA: Palmer School of Chiropractic. 1934 (1986 printing): 500, 508-510, 863-870.
8. Association of Chiropractic Colleges (ACC) Position on Chiropractic; Position Paper #1; July 1996; <http://lifenet.life.edu/Other/acc.html>
9. Idler EL, Kasl S. Health perceptions and survival: Do global evaluations of health status really predict mortality? J. Gerontology: Social Sciences 1991; 46(2)S55-S65.
10. Idler EL, Benyamini Y. Self-rated health and mortality: A review of twenty-seven community studies. J. Health Soc. Behav. 1997; 38:21-37.
11. Guyatt GH, Feeny DH, Patrick DL. Measuring Health-related Quality of Life. Ann. Int. Med. 1993; 118:622-629.
12. Woodruff SI, and Conway TL. Longitudinal assessment of the impact of health/fitness status and health behavior on perceived quality of life, Percept. Mot. Skills 1992; 5:3-14.
13. Sudman S. Applied Sampling. In: Rossi PH, Wright JD, Anderson AB. eds. Handbook of Survey Research. New York: Academic Press, 1983: 145-194.
14. Andrews FM, Withey SB. Social indicators of well-being: Americans' perceptions of life quality. New York: Plenum, 1976.
15. Caplan RD, Abbey A, Abramis DJ, et al. Tranquilizer use and well-being: a longitudinal study of social and psychological effects. Ann Arbor, MI: Institute for Social Research, 1984.
16. Spector PE. Summated Rating Scale Construction. Newbury Park, CA: Sage Publications. 1992.
17. Spitzer RL, Williams JBW, Gibbon M, et al. Structured Clinical Interview for DSM-III-R - Non-Patient Edition (SCID-NP, Version 1.0), Washington, DC. American Psychiatric Press, 1990.
18. Ware JE, Snow KK, Kosinski M, et al. SF-36 Health survey: manual and interpretation guide. Boston: The Health Institute, New England Medical Center; 1993.
19. Carmines EG, Zeller RA. Reliability and Validity Assessment. Newbury Park CA: Sage Publications, 1979.
20. Kim J, Mueller CW. Factor Analysis: Statistical Methods and Practical Issues. Newbury Park, CA: Sage Publications. 1978.
21. Nunnally JC. Psychometric theory 2nd ed. New York: McGraw-Hill, 1978.
22. Hawke C, Morter MT Jr. The use of measures of general health status in chiropractic patients: a pilot study. Palmer Journal of Research 1995; 2(2):39-45.
23. Snedecor GW, Cochran WG. Statistical Methods 7th ed. Iowa: Iowa State University Press, 1980.
24. Kazis LE, Anderson JJ, Meenan RF. Effect sizes for interpreting changes in health status. Med Care 1989; 27(3):S178-S189.
25. Jaeschke R, Singer J, Guyatt G. Measurement of health status. Ascertaining the minimal clinically important difference. Controlled Clin Trials 1989: 10:407-415.
26. Wolf FM. Meta-Analysis: Qualitative methods for research synthesis. Beverly Hills, CA: Sage Publications, 1986.
27. Cohen J. Statistical power analysis for the behavioral sciences. New York: Academic Press, 1977:8.
28. Iversen GR, Norpoth H. Analysis of Variance, 2nd Ed. Newbury Park, CA: Sage Publications, 1987.
29. Peters RK, Benson H, Porter D. Daily relaxation response breaks in a working population: Effects on self-reported measures of health, performance, and well-being, Am. J. Public Health 1977; 67:946-953.
30. Benson H, Goodale IL. The relaxation response: Your inborn capacity to counteract the harmful effects of stress, J. Fla. Med. Assoc. 1981; 68:265-267.
31. Sachar EJ, Fishman JR, Mason JW. Influence of the hypnotic trance on plasma 17-hydroxycorticosteroid concentration, Psychosom. Med. 1965; 27:330-341.
32. Michaels RR, Huber MJ, McCann DS. Evaluation of Transcendental Meditation as a method of reducing stress, Science 1976; 192:1242-1244.
33. Han JS, Terenius L. Neurochemical basis of acupuncture analgesia, Ann. Rev. Pharmacol. Toxicol. 1982: 22:193-220.
34. Steptoe A. The Assessment of sympathetic nervous function in human stress research, J. Psychosom. 1987; 31:141-152.
35. Jin P. Changes in heart rate, noradrenaline, cortisol and mood during Tai Chi. J. Psychosom. Res. 1989; 33:197-206.
36. Jin P. Efficacy of Tai Chi, brisk walking, meditation, and reading in reducing mental and emotional stress. J. Psychosom. Res. 1992; 36:361-370.
37. Heber L. Dance Movement: A therapeutic program for psychiatric clients. Perspect. Psychiatr. Care 1993; 29:22-29.
38. Littman AB, Fava M, Halperin P, et al. Physiological benefits of a stress reduction program for healthy middle-aged army officers., J. Psychosom.. 1993: 37:345-354.
39. Babbie E. The Practice of Social Research. 6th ed. Belmont, CA: Wadsworth. 1992.
40. U.S. Census Bureau; "United States population estimates, by age, sex, race, and hispanic origin, 1990 to 1996;" <http://www.census.gov/ >; (accessed: 23 June 1997).
41. Eisenberg DM, Kessler RC, Foster C, et al. Unconventional medicine in the United States; prevalence, costs, and patterns of use. New Engl. J. of Med. 1993; 328(4):246-252.

42. Fulder SJ, Munro RE. Complementary medicine in the United Kingdom: patients, practitioners, and consultations. The Lancet Sept. 1985; (2)542-545.
43. Himmel W, Schulte M, Kochen MM. Complementary medicine: are patients' expectations being met by their general practitioners? British J. Gen. Prac. June 1993; (43)232-235.
44. Shekelle PG, Markovich M, Louie R. Factors associated with choosing a chiropractor for episodes of back pain care. Medical Care 1995; (33)8:842-850.
45. Hansen JP, Futch DB. Chiropractic services in a staff model HMO: utilization and satisfaction. HMO Practice, 1997; 11(1):39-42.
46. Vladeck' BC. Equity, access, and the costs of health services. In: Securing Access to Health Care, vol. 3, President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research. Washington, D.C.: U.S. Government Printing Office, 1983: 9.
47. Mutchler JE, Burr JA. Racial differences in health and health care service utilization in later life: The effect of socioeconomic status. J. Health & Soc. Behav., 1991; 32:342-356.
48. Suchman EA. Social patterns of illness and medical care. J. Health & Soc. Behav., 1965; 6:2-16.
49. Geertsen R, Klauber MR, Rindflesh M, et al. An re-examination of Suchman's views on social factors in health care utilization. J. Health & Soc. Behav., 1975; 16:226-237.
50. Cleary PD, Mechanic D, Greenley JR. Sex differences in medical care utilization: An empirical investigation. J. Health & Soc. Behav. 1982; 23:106-119.
51. Verbrugge LM. The Twain meet: Empirical explanations of sex differences in health and mortality. J. Health Soc. Behav. 1989; 30:282-304.
52. Woodruff SI, Conway TL. Perceived quality of life and health-related correlates among men aboard Navy ships. Military Psych. 1990; 2:79-94.

APPENDIX

Self-Rated Health Scales: The following domains and items were used to assess health, wellness and overall quality of life.

I. Physical State

"Rate the following questions on a frequency scale of 1 to 5, with 1 = never, 2 = rarely, 3 = occasionally, 4 = regularly, 5 = constantly."

1. Presence of physical pain (neck/back ache, sore arms/legs etc.)
2. Feeling of tension or stiffness or lack of flexibility in your spine.
3. Incidence of fatigue or low energy.
4. Incidence of colds and flu.
5. Incidence of headaches (of any kind).
6. Incidence of nausea or constipation.
7. Incidence of menstrual discomfort.
8. Incidence of allergies or eczema or skin rashes.
9. Incidence of dizziness or lightheadedness.
10. Incidence of accidents or near accidents or falling or tripping.

II. Mental/Emotional State

"Rate the following questions on a frequency scale of 1-5, with 1 = never, 2 = rarely, 3 = occasionally, 4 = regularly, 5 = constantly."

1. If pain is present, how distressed are you about it.
2. Presence of negative or critical feelings about yourself.
3. Experience of moodiness or temper or angry outbursts.
4. Experience of depression or lack of interest.
5. Being overly worried about small things.
6. Difficulty thinking or concentrating or indecisiveness.
7. Experience of vague fears or anxiety.
8. Being fidgety or restless; difficulty sitting still.
9. Difficulty falling or staying asleep.
10. Experience of recurring thoughts or dreams.

III. Stress Evaluation

"Evaluate your stress relative to the following, with 1 = none, 2 = slight, 3 = moderate, 4 = pronounced, and 5 = extensive."

1. Family.
2. Significant Relationship.
3. Health.
4. Finances.
5. Sex Life.
6. Work.
7. School.
8. General well-being.
9. Emotional well-being.
10. Coping with daily problems.

IV. Life Enjoyment

"Rate the following questions on a degree scale of 1-5, with 1 = not at all, 2 = slight, 3 = moderate, 4 = considerable, 5 = extensive."

1. Openness to guidance by your "inner voice/feelings."
2. Experience of relaxation or ease or well-being.
3. Presence of positive feelings about yourself.
4. Interest in maintaining a healthy lifestyle (e.g., diet, fitness, etc.).
5. Feeling of being open and aware/connected when relating to others.
6. Level of confidence in your ability to deal with adversity.
7. Level of compassion for, and acceptance of, others.
8. Satisfaction with the level of recreation in your life.
9. Incidence of feelings of joy and or happiness.
10. Level of satisfaction with your sex life.
11. Time devoted to things you enjoy.

Overall Quality of Life (Woodruff and Conway, 1992)

"Evaluate your feelings relative to the quality of your life with 1 = terrible, 2 = unhappy, 3 = mostly dissatisfied, 4 = mixed, 5 = mostly satisfied, 6 = pleased, 7 = delighted."

1. Your personal life.
2. Your wife/husband or (significant other).
3. Your romantic life.
4. Your job.
5. Your co-workers.
6. The actual work you do.
7. Your handling of problems in your life.
8. What you are actually accomplishing in your life.
9. Your physical appearance - the way you look to others.
10. Your self.
11. The extent to which you can adjust to changes in your life.
12. Your life as a whole.
13. Overall contentment with your life.
14. The extent to which your life has been what you wanted it to be.

Clinton Presidential Records
Digital Records Marker

This is not a presidential record. This is used as an administrative marker by the William J. Clinton Presidential Library Staff.

This marker identifies the place of a tabbed divider. Given our digitization capabilities, we are sometimes unable to adequately scan such dividers. The title from the original document is indicated below.

Divider Title: Hammerschlag

Richard Hammerschlag, Ph.D.
Research Director, Oregon College of Oriental Medicine
10525 Cherry Blossom Drive
Portland, OR 97216
503/253-3443 x156
rhammerschlag@ocom.edu

A doctorate in biochemistry from Brandeis University, Waltham, Massachusetts, in 1967 launched Richard Hammerschlag's 25-year career in neuroscience research, based mainly at the City of Hope National Medical Center, Duarte, California. While there, he served as associate chair of the Division of Neurosciences from 1986 to 1995 and published over 50 research articles and scientific book chapters.

Earlier, in the late 1970's, a personal experience with acupuncture triggered what would become Hammerschlag's ongoing interest in exploring this traditional medical practice in terms of modern Western medicine: his acute sciatica condition was successfully treated by a doctor of Oriental medicine. In 1991, he was invited to create and teach a course in "Biomedical Understanding of Acupuncture" at Yo San University of Traditional Chinese Medicine, Santa Monica, California. Several years of teaching the course culminated in a major career change when, in 1995, he accepted the positions of academic dean and research director at the same university. In the fall of 1999, he became research director at the Oregon College of Oriental Medicine, a move that also allowed him to live near his two young grandchildren.

His main research interests are:

- Evaluating the clinical efficacy of acupuncture and Chinese herbal therapy using research designs that most closely reflect the real-world clinical practice of this traditional medical system.
- Examining what acupuncture in specific and CAM in general can tell us about how the body works that Western biomedicine has not yet discovered.

Hammerschlag's national-level acupuncture activities include:

- 1993: Member, initial grants review panel, Office of Alternative Medicine, NIH
- 1994: Chair, methodology panel, FDA conference on the status of the acupuncture needle
- 1996: Co-author, with Stephen Birch, *Acupuncture Efficacy: A Compendium of Controlled Clinical Trials*, National Academy of Acupuncture and Oriental Medicine, Tarrytown, NY.
- 1997: Invited speaker on "Methodological and Ethical Issues in Acupuncture Research", NIH Consensus Development Conference on Acupuncture
- 1997 to present: Co-President, Society for Acupuncture Research
- 1999 to present: Co-Investigator of clinical trials of acupuncture at the two NCCAM-funded, Portland-based CAM research centers
- 2001: Co-editor, with Gabriel Stux, *Clinical Acupuncture: Scientific Basis*, Springer-Verlag, Berlin.

Richard Hammerschlag, PhD
Research Director
Oregon College of Oriental Medicine
10525 SE Cherry Blossom Drive
Portland, OR 97216
503/253-3443, x156
rhammerschlag@ocom.edu

Testimony to the
White House Commission on Complementary and Alternative Medicine Policy
May 16, 2001

Good morning, and thank you for this opportunity to talk with you on issues related to CAM research training. I am fortunate to have had two careers in research: one spanning 25 years at the biomedical lab bench, the other, now reaching 10 years, immersed in the challenges of exploring acupuncture and CAM. At present I am the research director at the Oregon College of Acupuncture and Oriental Medicine, involved in CAM research at each of the two NCCAM-funded centers in Portland, and I serve as co-president of the Society for Acupuncture Research. From my dual perspectives of biomedical and CAM research, it is a pleasure to offer input for the strategies you are developing to help sail the ship of CAM on the seas of evidence-based medicine.

I would like to address three concerns:

1. The need for CAM *research* more clearly to reflect CAM *practice*.
2. The need for better training of both CAM and conventional practitioners in the skills necessary to evaluate the quality of CAM research articles.
3. The need to examine systematically what CAM practice and research are revealing about how the body works that biomedicine has not yet figured out.

First, I will address why I feel that much of CAM research has been inadequate in its evaluation of CAM practice. The need for CAM research, as reflected in the

#40

creation of OAM and its upgrade to NCCAM, has been recognized because of the widespread public use of a wide range of under-researched CAM therapies. It follows that what needs to be evaluated in clinical trials is how people are being treated. In contrast, what too often appears in clinical trials is a CAM therapy morphed into an allopathic-like treatment resulting in standardized protocols and variably designed, non-validated placebo treatments. The challenge for CAM research is to maintain rigorous research standards while having the protocol reflect the real-world practice of the CAM therapies.

When CAM practitioners read a research study, their response should not be, "This research protocol is far too restrictive or simplified. I would never treat a patient that way." Rather, their response should be "Well, I might not treat my patient in exactly that way, but the protocol is reasonable, given the presenting condition and diagnosis."

While my years in biomedical research have left me with an appreciation of the need for controlled trials, I believe, given the limited resources available for funding CAM research, that undue emphasis is placed on placebo-controlled trials. What I see a greater need for are what researchers call trials of "pragmatic efficacy", in which two treatments are compared under conditions in which they would be applied in practice. As Kaptchuk, Edwards and Eisenberg have recently written (1996, p.44), "A pragmatic perspective provides less scientifically useful information but potentially provides more clinically relevant information." As an example, a patient with tennis elbow -- and the patient's acupuncturist -- are less interested in a clinical trial that has demonstrated acupuncture as more effective than placebo acupuncture than in an equally rigorous trial in which acupuncture was found to be as effective as steroid injections, and with fewer adverse effects (MacPherson, 1999). Even when a trial compares treatment to placebo, the treatment should reflect real-world practice.

An excellent example is the recent trial of Chinese herbal therapy for Irritable Bowel Syndrome (Bensoussan et al, 1998). A group of patients receiving placebo herbs was compared to a group receiving a standardized herbal formula for IBS and to a third group receiving herbs individually formulated to each patient's Chinese medicine diagnosis.

As a further consideration, not only should acupuncture or Chinese herbal therapy be evaluated in the manner each is delivered in the clinic, they also need to be evaluated as an integrated whole, as they are delivered within the system of Traditional Chinese Medicine. While I am not aware of any published study in which the effectiveness of TCM has been evaluated as a *system* of care, I am pleased to be involved in such a study at the present time. At the NCCAM Center based at Kaiser-Permanente's Center for Health Research, we are currently comparing the systems of TCM, Naturopathic Medicine and Allopathic specialty care for women with complex temporomandibular disorders (TMD).

Recommendations:

1. That NCCAM be encouraged to call for research proposals that clearly reflect the real-world practice of CAM.
2. That CAM research proposals submitted to NCCAM and to the various NIH Institutes be required to have had CAM practitioners involved in the research design.
3. That the CAM research design in such proposals must be accompanied by a clear rationale and must have been approved by a consensus panel of experienced CAM practitioners.
4. That representatives from the NCCAM-funded Centers should meet periodically not only to present progress reports on their research but to exchange experiences of how best to educate researchers in CAM practices and how best to introduce CAM practitioners to the research culture. The

NCCAM Centers should be strongly encouraged to strengthen their collaborations between experienced clinical researchers and CAM practitioners.

I will be briefer in presenting the second and third of my three general concerns. My second concern stems from teaching and lecturing about CAM research in general and acupuncture research in particular to students and practitioners of CAM as well as conventional medicine. What I repeatedly encounter, in this era of "evidence-based medicine", is a serious need for training (and for continuing education) in how to understand and evaluate CAM clinical research studies. Most practitioners, CAM or conventional, are neither well trained in the general principles of research design nor appreciate the difficulties inherent in applying conventional clinical research models to CAM research. *Yet increasingly in their practice, they are facing the task of reading and evaluating CAM research trials.* Too often there is more attention paid to whether a trial's outcome is positive or negative rather than to whether the trial was designed well-enough to have provided an adequate test of the CAM therapy.

In this context, Systematic Reviews are helpful but of limited value. As you are aware, Systematic Reviews apply criteria of research quality to selected trials of a particular therapy for a specific condition. Such reviews were developed to be able to draw general conclusions of the research-based efficacy of conventional treatments, but the same approach has been increasingly applied to CAM research. I believe the conclusions of these reviews are of limited value for CAM since it is relatively rare to find a sufficient number of similar trials to allow a proper Systematic Review to be performed. The reality is that a surprisingly wide variety of procedures, all referred to as acupuncture (similar for massage or chiropractic) are pooled together for review. A Systematic Review of acupuncture typically includes trials of manual- as well as electro-acupuncture,

trials in which widely variable numbers of needles were inserted into acupuncture points, trigger points, tender points or muscle motor points, and trials in which acupuncture was compared to a broad range of invasive and non-invasive placebo treatments or to a broad range of standard care procedures. (Ezzo et al, 2001). A second reason why such reviews are of limited value is that they rarely attempt to assess whether the treatment protocols were adequate for the given condition.

Again, it has been my experience that without an appreciation of the inherent difficulties in designing and performing CAM research, practitioners of CAM and conventional medicine alike are not able to adequately assess the conclusions of these Systematic Reviews.

Recommendations:

1. That students at allopathic medical schools (and CAM colleges) must be taught how to understand and critique CAM research trials. It is necessary but not sufficient to offer survey courses of CAM practices, as an increasing number of allopathic medical schools have been doing; *there must also be survey courses in CAM research.*
2. That CAM and conventional medicine practitioners must take continuing education seminars in how to understand and critique CAM research trials.

My first two concerns have focused on the question "Does CAM work?" I stressed the need for CAM research to be better informed by and more relevant to CAM practice. And, I emphasized the need for better training of both CAM and conventional practitioners in the critical skills necessary to evaluate the quality of CAM research.

My final concern addresses the question, "How does CAM work?"

With all the current emphasis on clinical trials, we are not paying enough attention to a profoundly interesting parallel development of CAM practice and research: CAM is revealing ways in which the body works that biomedicine has not yet figured out. We need to begin a concerted effort to expand our concepts of physiology to incorporate many of the novel features of CAM practice that are being validated by CAM research. We hear and read of tentative proposals defining a new field of Energy Medicine that may bring together a variety of traditional healing practices with state of the art medical technology (Rubik, 1995; Oschman, 2000). We must complement these creative efforts by seeking to develop an Energy Physiology that will serve as the foundation of this emerging view of the body.

As we set off in these directions, there are three broad areas to explore:

- The psycho-physiology of the placebo response.
- The physiology underlying Qigong, Reiki, Therapeutic Touch, and other CAM practices that are defining the landscape and boundaries of Energy Medicine.
- The basis of healing practices, including distant healing and intercessory prayer that, unlike the phenomena underlying energy practices, appear to be non-mediated as well as non-local.

While research on the *physiology* of placebo effects is certainly not solely within the province of CAM, biomedicine appears little interested. Yet an understanding of the essential feature of the placebo response: *the transformation of meaning into molecules*, will surely provide us with major insights into a range of mind-body CAM phenomena, as well as improve the non-specific aspects of

healing. As Harris Dienstfry elegantly understates, "The mind as a source of medicine is waiting to be explored." (Dienstfry, 1999)

Second, while the term energy medicine is often used today in an amorphous manner, research on bioelectromagnetic fields (so-called BEM fields) produced by every organ and tissue of the body, and detected at the palms of Qigong masters (Seto et al, 1992), is giving form and substance to this term. Even at the molecular level, there is a sea change in thinking, in which biochemistry is being considered largely as a solid-state phenomenon (Ho, Knight, 1998). In this view enzyme-mediated activity as well as intra- and inter-cellular signaling occur on and along cytoskeletal structures and proteins that span cell membranes. The fixed ionic charges at the surface of these structures allow electromagnetic signaling within and between cells and tissues at rates orders of magnitude faster than signals mediated by diffusible biochemicals.

An understanding of the nature and origin of BEM fields, and how they may be modulated by external means, has great potential value for human health. This may well serve as a modern-day approach to the goal of many traditional systems of healing: The detection *and correction* of imbalances in the body before they reach the stage of clinical signs and symptoms recognized by allopathic medicine.

Lastly, we must expand our understanding of the human body by the scientific study of healing phenomena that do not appear explainable in terms of either molecular or energy medicine. The marked lessening of symptoms in HIV-positive patients (Sicher et al, 1998) and the physiological improvement of patients in coronary care units (Harris et al, 1999) are recent examples of randomized controlled trials assessing a therapeutic practice involving mental intention by healers at sites remote from patients. How is such healing

"received" by patients and how does it alter physiology? Answers to such questions are of profound importance for redefining our concepts of self and other, and for the reshaping of future healthcare.

Recommendations:

1. That medical school survey courses of CAM and Integrative Medicine should include research findings and emerging concepts that challenge the prevailing views of physiology and pathophysiology.
2. That NCCAM -- or perhaps the National Academy of Sciences -- should be encouraged to call for a conference that would propose evidence-based models for expanding our present concepts of physiology, and that would recommend an agenda for future research.

REFERENCES

Bensoussan A, Talley NJ, Hing M, Menzies R, Guo A, Ngu M (1998) Treatment of irritable bowel syndrome with Chinese herbal medicine: A randomized controlled trial, *JAMA* 280:1585-1589.

Dienstfrey H (1999) Mind and mindlessness in mind-body research, *Advances in Mind-Body Med* 15:229-233.

Ezzo J, Lao L, Berman BM (2001) Assessing clinical efficacy of acupuncture: What has been learned from systematic reviews of acupuncture? In: *Clinical Acupuncture: Scientific Basis* (G Stux, R Hammerschlag, eds), Springer-Verlag, Berlin.

Harris WS, Gowda M, Kolb JW, Strychacz CP, Vacek JL et al (1999) A randomized, controlled trial of the effects of remote, intercessory prayer on outcomes in patients admitted to the coronary care unit, *Arch Intern Med* 159:2273-2278.

Ho MW, Knight DP (1998) The acupuncture system and the liquid crystalline Collagen fibers of the connective tissues, *Am J Chin Med* 26:251-263.

Kaptchuk TJ, Edwards RA, Eisenberg DM (1996) Complementary medicine: efficacy beyond the placebo effect, In: *Complementary Medicine: an objective appraisal* (Ernst E, ed.) Butterworth/Heinemann, Oxford, pp. 42-70.

MacPherson H (1999) Out of the laboratory and into the clinic: acupuncture research in the real world, *Clin Acupunct Oriental Med* 1:97-100.

Oschman JL (2000) *Energy Medicine: The Scientific Basis*, Churchill Livingstone, Edinburgh

Rubik B (1995) Energy medicine and the unifying concept of information, *Altern Ther* 1:34-39.

Seto A, Kusaka C, Nakazato S, Huang WR, Sato R et al (1992) Detection of extraordinary large biomagnetic field strength from human hand during external qi emission, *Acupunct Electrother Res* 17:75-94.

Sicher F, Targ E, Moore D, Smith HS (1998) A randomized double-blind study of the effect of distant healing in a population with advanced AIDS: Report of a small-scale study, *West J Med* 169:356-363.

White House Commission on Complementary and Alternative Medicine:

Publication of Research Results

A fundamental obligation of every type of medicine is to identify what treatments can help which patients and describe the basis of that information in a way that is objective and useful. Researchers must identify the specific questions that they are attempting to answer and undertake their studies recognizing that their beliefs and prior impressions may turn out to be mistaken. For all types of practices related to health it is useful to gather information about what, in fact, current practices, beliefs, and economic consequences are. This is particularly so with unconventional medicine because physicians may be unaware of what care patients seek outside the medical establishment. In clinical studies a constant challenge is to identify designs and conduct the research in ways that minimize bias, including biases that may come from either skeptics or enthusiasts, from investigators or patients. In clinical investigations comparison with appropriate controls will increase confidence in the results as will comparisons with various types of treatments that are currently in use. Research evaluations should try to identify what degrees of change are clinically meaningful, recognizing that small differences, even if real, may not be of any real benefit to patients. Evaluations should try to assess outcomes in many ways, including effects on patients' function, physiological changes, and changes in symptoms, as well as possible adverse effects from any intervention. Also relevant are the costs of interventions and any effects on the subsequent use of health services. In all clinical studies, longer, more complete follow-up increases the strength of the evidence.

The first commitment of medical-scientific journal is to publish research that is accurate and relevant to the journal's readers. Hence, the journals must assess both the evidence in a report and the process by which it was obtained. Some of the questions that editors ask are: how important is this research? What exactly is the question trying to be answered? How rigorously was the work conducted? Are the results reproducible, generalizable? Were the data analyzed in a way that is valid statistically? Finally, is the work presented in a way that is clear and comprehensible and that places it in context of other research that has been done on this topic?

-- Edward W. Campion, M.D.

Clinton Presidential Records
Digital Records Marker

This is not a presidential record. This is used as an administrative marker by the William J. Clinton Presidential Library Staff.

This marker identifies the place of a tabbed divider. Given our digitization capabilities, we are sometimes unable to adequately scan such dividers. The title from the original document is indicated below.

Divider Title: Peer Reviewed Journals

VII

Questions for the Panel on Peer-Reviewed Journals:

What criteria are used by your journal for accepting articles on CAM research?

How do you choose your reviewers; what qualifications must they have? Are there standards for judging reviewer expertise?

Are the criteria used by conventional journals in accepting CAM studies similar to or different from criteria used for acceptance of conventional studies? Is there any outreach by conventional journals to include studies on CAM interventions?

What are the unique challenges facing CAM acceptance in conventional peer-reviewed journals?

Are the criteria used by CAM journals in accepting research studies comparable to the standards used by conventional journals?

What are the unique challenges facing CAM peer-reviewed journals?

What are the most common reasons articles are rejected by your journal? Are you able to calculate the rate of rejected submissions in general; for CAM in particular?

On average, how many revisions are made before publication?

Does your audience include both CAM practitioners; conventional practitioners? What percentage of your readership do you believe falls into either category?

What policy recommendations can you offer regarding the publication of CAM practices and products in peer-reviewed journals for the Commission's consideration as it develops its report to the President and Congress?

Dr. Edward W. Campion has been a Deputy Editor at the *New England Journal of Medicine* since 1987 where he has written editorials on unconventional medicine, the Internet, and the doctor-patient relationship. Dr. Campion is certified in internal medicine and geriatric medicine. He is a graduate of Dartmouth Medical School, and completed clinical training at Cambridge Hospital and the Massachusetts General Hospital where he was the Chief of the Geriatrics Unit from 1979 until 1987.

White House Commission on Complementary and Alternative Medicine:

Publication of Research Results

A fundamental obligation of every type of medicine is to identify what treatments can help which patients and describe the basis of that information in a way that is objective and useful. Researchers must identify the specific questions that they are attempting to answer and undertake their studies recognizing that their beliefs and prior impressions may turn out to be mistaken. For all types of practices related to health it is useful to gather information about what, in fact, current practices, beliefs, and economic consequences are. This is particularly so with unconventional medicine because physicians may be unaware of what care patients seek outside the medical establishment. In clinical studies a constant challenge is to identify designs and conduct the research in ways that minimize bias, including biases that may come from either skeptics or enthusiasts, from investigators or patients. In clinical investigations comparison with appropriate controls will increase confidence in the results as will comparisons with various types of treatments that are currently in use. Research evaluations should try to identify what degrees of change are clinically meaningful, recognizing that small differences, even if real, may not be of any real benefit to patients. Evaluations should try to assess outcomes in many ways, including effects on patients' function, physiological changes, and changes in symptoms, as well as possible adverse effects from any intervention. Also relevant are the costs of interventions and any effects on the subsequent use of health services. In all clinical studies, longer, more complete follow-up increases the strength of the evidence.

The first commitment of medical-scientific journal is to publish research that is accurate and relevant to the journal's readers. Hence, the journals must assess both the evidence in a report and the process by which it was obtained. Some of the questions that editors ask are: how important is this research? What exactly is the question trying to be answered? How rigorously was the work conducted? Are the results reproducible, generalizable? Were the data analyzed in a way that is valid statistically? Finally, is the work presented in a way that is clear and comprehensible and that places it in context of other research that has been done on this topic?

-- Edward W. Campion, M.D.

Clinton Presidential Records
Digital Records Marker

This is not a presidential record. This is used as an administrative marker by the William J. Clinton Presidential Library Staff.

This marker identifies the place of a tabbed divider. Given our digitization capabilities, we are sometimes unable to adequately scan such dividers. The title from the original document is indicated below.

Campion

Divider Title: _____

Dr. Edward W. Campion has been a Deputy Editor at the *New England Journal of Medicine* since 1987 where he has written editorials on unconventional medicine, the Internet, and the doctor-patient relationship. Dr. Campion is certified in internal medicine and geriatric medicine. He is a graduate of Dartmouth Medical School, and completed clinical training at Cambridge Hospital and the Massachusetts General Hospital where he was the Chief of the Geriatrics Unit from 1979 until 1987.

White House Commission on Complementary and Alternative Medicine:

Publication of Research Results

A fundamental obligation of every type of medicine is to identify what treatments can help which patients and describe the basis of that information in a way that is objective and useful. Researchers must identify the specific questions that they are attempting to answer and undertake their studies recognizing that their beliefs and prior impressions may turn out to be mistaken. For all types of practices related to health it is useful to gather information about what, in fact, current practices, beliefs, and economic consequences are. This is particularly so with unconventional medicine because physicians may be unaware of what care patients seek outside the medical establishment. In clinical studies a constant challenge is to identify designs and conduct the research in ways that minimize bias, including biases that may come from either skeptics or enthusiasts, from investigators or patients. In clinical investigations comparison with appropriate controls will increase confidence in the results as will comparisons with various types of treatments that are currently in use. Research evaluations should try to identify what degrees of change are clinically meaningful, recognizing that small differences, even if real, may not be of any real benefit to patients. Evaluations should try to assess outcomes in many ways, including effects on patients' function, physiological changes, and changes in symptoms, as well as possible adverse effects from any intervention. Also relevant are the costs of interventions and any effects on the subsequent use of health services. In all clinical studies, longer, more complete follow-up increases the strength of the evidence.

The first commitment of medical-scientific journal is to publish research that is accurate and relevant to the journal's readers. Hence, the journals must assess both the evidence in a report and the process by which it was obtained. Some of the questions that editors ask are: how important is this research? What exactly is the question trying to be answered? How rigorously was the work conducted? Are the results reproducible, generalizable? Were the data analyzed in a way that is valid statistically? Finally, is the work presented in a way that is clear and comprehensible and that places it in context of other research that has been done on this topic?

-- Edward W. Campion, M.D.

Clinton Presidential Records
Digital Records Marker

This is not a presidential record. This is used as an administrative marker by the William J. Clinton Presidential Library Staff.

This marker identifies the place of a tabbed divider. Given our digitization capabilities, we are sometimes unable to adequately scan such dividers. The title from the original document is indicated below.

Wootton

Divider Title: _____

Jackie Wootton, M.Ed.

Alternative Medicine Foundation, Inc.
5411 West Cedar Lane, Suite 205-A, Bethesda, MD 20814, USA
Tel: (301) 581-0116; Fax: (301) 581-0119; email: jcw@amfoundation.org
http://www.amfoundation.org

Brief biographical details:

Jackie Wootton, M.Ed. has been Executive Editor of the Journal of Alternative & Complementary Medicine, published by M A Liebert, Inc, New York, since 1996. The editorial office is run from the Alternative Medicine Foundation, Inc which Ms. Wootton founded in 1998, of which she is President and Executive Director. One of the main programs she initiated for the foundation is the internationally recognized HerbMed® herbal database.

As an Associate Investigator at the NIH Clinical Center, she designed three questionnaires for surveys of CAM use by clinical center patients. These were published in 2000, as well as a joint book chapter on CAM surveys. She is currently jointly publishing a series of articles on different categories of CAM surveys. Ms. Wootton has published widely on alternative medicine research information, women's health research, electronic databases, and the Internet, including regular WebWatch columns for health and clinical research journals.

Educated in the UK, Ms. Wootton became a professor of sociology at Leeds Metropolitan University, UK (1985-1992), Course Director for Consumer Services Management (1990-1992), specializing in methods of social research, survey design and analysis; cultural, consumer and lifestyles studies. Concurrently (1980-1990), she worked as a part time tutor, counselor, course developer, and staff trainer for 10 years for the Open University, UK, the pioneer distance learning institution. On moving to the USA , she worked for 2 years (1993-1995) designing and developing information resources for the Office of Alternative Medicine, NIH, followed by an independent consultancy as Informatics Project Director at the Rosenthal Center for Complementary and Alternative Medicine at Columbia University, College of Physicians & Surgeons, NY (1995-2000).

SUMMARY OF REMARKS FOR WHITE HOUSE COMMISSION MEETING ON PEER-REVIEWED JOURNALS ON MAY 16, 2001

PANELIST: Jackie Wootton, M.Ed., Executive Editor of the *Journal of Alternative and Complementary Medicine*, published by Mary Ann Liebert, Inc, New York.

We believe that the review process for the *Journal of Alternative and Complementary Medicine* is essentially the same as that of any major biomedical research journal. We practice the same high standards of expert and impartial peer review that are accepted practice for all quality academic publications.

The journal was accepted for indexing by Medline within its first three years of publication, backdated to the start year and has steadily gained a strong reputation as a serious CAM research publication. When the journal changed from quarterly to bi-monthly in 1999, more libraries subscribed. Currently, about 60% of our subscriptions are Medical or Biomedical School libraries and libraries in research institutions. Individual subscribers are equally divided between clinicians and CAM practitioners, researchers, and others (administrators, medical societies, people in the natural products industries, other business subscribers). For the most part, the lay public does not read the journal.

When selecting reviewers, the authors' recommendations and requests are taken into consideration, also requests for excluding certain reviewers. We have a large editorial board who review and are available for consultation and an extensive database of expert reviewers. Review is anonymous (but not double-blinded). There are usually 3 reviewers: one with the appropriate CAM knowledge and expertise; one conventional biomedical researcher, one methodologist or statistician, although this requirement can be difficult to sustain.

The criteria for accepting articles is that they are clear and coherent and scholarly, that description of methods and materials is complete and replicable, that studies are contextualized within the appropriate existing bodies of knowledge, and that claims and conclusions are supported by the evidence.

Like all the major journals, we receive promotional articles sponsored by companies who need "third party literature". Like all the major journals we request a declaration of interest when a manufacturing company is involved. As for all the major conventional medicine journals, occasional weak papers slip through. No peer review process is infallible.

The main reasons for rejection of papers is:

- Weak, unsubstantiated claims
- Strong claims lacking adequately strong support
- Trivial, overblown papers
- Promotional papers, and
- Lack of rigor and objectivity

These problems face all journal editors. What I see as **the unique challenges** facing CAM journals, and our journal specifically, are:

- Paucity of material - I receive approximately 110 articles a year
- Many researchers have not been exposed to stringent scientific standards before
- Lack of experienced researchers who have CAM experience and insights
- Lack of reviewers who understand the code of expert and impartial review
- Non-English speaking authors have difficulty expressing ideas clearly

From unique challenges come **unique opportunities**. Editors spend a lot of time educating and encouraging authors. Rather than reject out of hand we help authors to modify their claims. For example, we may require them to add statements on the limitations of their experiment or methodology, or to rework a paper. For instance, a poorly designed clinical trial may have potentially interesting preliminary data or insights that can be converted into a promising case study or a short report. We tolerate successive stages of review and revision. Most papers go through one major revision stage but some require several iterations. Our rejection rate is low, around 30%, as a result of our educative policy. The attrition rate is high; many authors do not return a revised paper for reconsideration.

Although the underlying principles of publication are the same for CAM as for conventional medical journals, the process does have differences. CAM articles that are published by conventional journals are predominantly based on the standard pharmaceutical drug model of isolating and testing single agents, typically using the randomized controlled trial. At the Journal of Alternative and Complementary Medicine we encourage a broader perspective, that new theories demand innovative methods. Provided researchers adhere to the basic scientific principles of external and internal validity, they should be free to innovate, develop new appropriate methodologies, and carefully crafted techniques.

Recommendations:

- for more academic departments and research positions at post-doctoral level to attract the "brightest and best" into the field
- for more research money for practitioner training schools as they are the ones with the knowledge and insights needed to develop appropriate methodologies
- to expand the range of acceptable models of quality clinical research beyond the formulaic, large clinical trial protocols - while retaining integrity and authority.

There is a particular need for sponsorship from independent funding sources that will encourage innovative and imaginative but rigorous research.

May 2, 2001

Clinton Presidential Records
Digital Records Marker

This is not a presidential record. This is used as an administrative marker by the William J. Clinton Presidential Library Staff.

This marker identifies the place of a tabbed divider. Given our digitization capabilities, we are sometimes unable to adequately scan such dividers. The title from the original document is indicated below.

Divider Title: Fontanarosa

Phil Fontanarosa, M.D.

Journal of the American Medical Association

#43

Biographical Information

Phil B. Fontanarosa, MD is Executive Deputy Editor of *JAMA, The Journal of the American Medical Association,* and is Vice-President of Scientific Publications and Multimedia Applications at the American Medical Association.

Dr. Fontanarosa has been an editor at *The Journal* since 1993. Dr. Fontanarosa is a graduate of the Medical College of Ohio, is board certified in emergency medicine, and also attended the Graduate School of Journalism at Kent State University.

Dr. Fontanarosa is an Adjunct Professor of Medicine at Northwestern University Medical School and is a member of the medical staff at Northwestern Memorial Hospital in Chicago, Illinois.

#43

Statement Before the White House Commission

on Complementary and Alternative Medicine Policy

Meeting on Research Challenges

May 16, 2001

"Publication of Peer-Reviewed CAM Research Results"

Phil B. Fontanarosa, MD

Executive Deputy Editor
Journal of the American Medical Association

Adjunct Professor of Medicine
Northwestern University Medical School
Chicago, Illinois

Dr. Gordon, members of the White House Commission, thank you for the opportunity to participate in this meeting and to provide information on the topic of publication of complementary and alternative medicine (CAM) research results in peer reviewed medical journals. My remarks are intended to briefly discuss three issues relevant to peer reviewed general medical journals: (1) Why should medical journals consider publishing research on complementary and alternative medicine? (2) How do peer reviewed medical journals evaluate research on complementary and alternative medicine? and, (3) What are some policy issues regarding publication of complementary and alternative medicine research in peer reviewed journals?

Why should medical journals consider publishing research on complementary and alternative medicine?

During the past 10 years, the use of complementary and alternative medicine and visits to alternative medicine practitioners have increased dramatically, along with marked increases in expenditures and increasingly prevalent third party reimbursement for these therapies. At academic medical centers, there has been increased interest among medical students and resident physicians in gaining knowledge about alternative therapies, and learning opportunities for CAM have been added to the medical education curriculum. In addition, there have been significant increases in federal funding for national centers dedicated to CAM research, along with increased funding from government and from industry to examine CAM therapies.

Yet, despite the increasing interest in and use of CAM, high quality scientific evidence that clearly establishes the safety and efficacy (or lack thereof) of these interventions is lacking. Consequently, many physicians traditionally have viewed CAM in general, and most practices contained therein, with skepticism and mistrust. Nonetheless, many patients use CAM therapies along with their conventional medical

therapies, and physicians should be the health professionals responsible for coordinating all medical care for their patients. Accordingly, physicians require reliable information on CAM to understand the complex interplay of CAM therapies and conventional therapies, to be prepared to serve as resources for their patients, to be comfortable answering their patients' questions about CAM, and to provide evidence-based guidance and advice about various CAM therapies. Ideally, this information should be based on critically evaluated, methodologically rigorous, and scientifically valid articles published in reputable journals, rather than on anecdotal reports, unfounded opinion, or unproven theories regarding CAM therapies.

To be relevant for physician readers, biomedical journals must respond to trends in medical science, react to trends among the public regarding medical therapies, and serve the needs of readers. In a 1997 survey, physician readers ranked alternative medicine as the seventh most important topic (of 70 choices) for *The Journal* to address, whereas the Editorial Board and editorial staff ranked alternative medicine in the top 3 subjects to address in the coming year. Thus, *JAMA* and the 9 other AMA journals, the *Archives* journals (*Archives of Internal Medicine, Archives of Pediatrics and Adolescent Medicine, Archives of General Psychiatry, Archives of Dermatology, Archives of Family Medicine, Archives of Neurology, Archives of Otolaryngology-Head and Neck Surgery, Archives of Ophthalmology, Archives of Surgery)* developed an initiative to objectively assess various CAM therapies and published coordinated theme issues devoted to CAM topics in November 1998. Through that initiative, more than 80 articles on CAM topics were published in the peer-reviewed biomedical literature in *JAMA* (9 research articles, including 6 randomized controlled trials, and 7 opinion articles) and the *Archives* Journals. Although I do not practice alternative medicine and have not conducted research in CAM therapies, I was responsible for planning,

coordinating, and editing this issue of *The Journal* on CAM research (*JAMA*. 1998;280:1549-1640), and also helped to coordinate efforts with the *Archives* Journals.

How do peer reviewed medical journals evaluate research on complementary and alternative medicine?

Peer reviewed general medical journals should be willing to consider papers on virtually any topic relevant to the practice of medicine, medical research, and their physician readers. Most research papers on CAM should not be rejected out-of-hand simply based on the topic, but should receive an unbiased, objective editorial evaluation, based on assessment of the following: the scientific importance of the study; the methodological rigor with which the research is conducted, the novel contribution of the research to the existing literature; and the relevance of the study results for the journal's readers. Evaluation of the scientific quality should involve the same rigorous and critical appraisal used to assess studies reporting research on conventional therapies, and the accepted framework of scientific evidence should prevail.

Some key aspects in the evaluation of CAM research studies include assessment of the following: an explicit, focused research question; scientific or biological plausibility; defined target disease or condition; established and accepted rigorous research methods (e.g., appropriate controls, effective blinding procedures, adequate power); measurable, objectively assessed endpoints (e.g., blinded assessment); meaningful, patient centered outcomes; information on adverse effects and safety data; and appropriate discussion of clinical context, public health importance, and study limitations.

The assessment of CAM research studies frequently involves peer review. Peer review of CAM studies involves selection of appropriate peer reviewers who have expertise in the research methods and subject matter of the study to obtain an objective

and unbiased evaluation of the research. Reviewers are identified and selected in several ways, but most commonly from a large database of *JAMA* reviewers (which includes self-designated areas of expertise, along with editors' ratings of reviewer quality) or from literature searches for individuals who have expertise with the research topic under evaluation. The peer reviewers serve as consultants who provide the editors with information regarding the content, context, and contribution of the study in light of existing research as well as critical appraisal of the research methods and statistical analyses. Peer reviewers also provide recommendations about the suitability of the study for publication and objective suggestions to improve the research report. Comments of peer reviewers are considered very carefully in editorial decision making about all research papers, including CAM research. Physician reviewers are now available for most topics evaluated in CAM research.

In general, decisions to publish CAM studies in a general medical journal like *JAMA* involve assessment of the quality of the research, the clinical applicability and importance of the research, and the generalizability of the research findings. An additional important consideration is the assessment of the merits of the study compared with the information provided in the many other papers submitted for consideration for publication. In large circulation general medical journals, competition for space is particularly keen, with a large number of submitted papers that span the entire gamut of medical research, clinical care, and medical and surgical specialties. In 2000, the overall acceptance rate for all manuscripts submitted to *JAMA* was approximately 10%, and for unsolicited manuscripts the acceptance rate was approximately 8%. With this competition for space, reports of large scale, high quality, multi-center, randomized trials on clinically important topics are more likely to receive priority for publication than single center studies reporting intermediate or preliminary results about specialized

CAM therapies. In fact, during the past year, studies on CAM therapies published in *JAMA* have been multi-center randomized trials; the most recent example is a multi-center randomized trial that showed lack of efficacy of St. John's Wort for treatment of depression (*JAMA*. 2001;285:1978-1986).

What are some policy issues regarding publication of CAM research in peer reviewed journals?

There are several possible policy issues regarding publication of CAM research in peer reviewed medical journals, and most of these were outlined in a previous editorial (*JAMA*. 1998;280:1618-1619)]. Carefully conducted, high quality, credible research by established biomedical investigators is necessary to evaluate the efficacy and safety of complementary and alternative medicine therapies. However, until convincing evidence is available that demonstrates efficacy, safety, and effectiveness of CAM interventions, uncritical acceptance and widespread application of these untested and unproven therapies must stop. With the recent increases in research funding and the establishment of national centers for CAM research, future publication of high quality studies on CAM therapies will establish the evidence base that supports or refutes the safety and efficacy of these therapies. Complementary and alternative medicine interventions that are shown to cause harm or that are demonstrated to have no beneficial effect should be abandoned. Decisions regarding use of and reimbursement for CAM therapies should be based on evidence based research findings and proper cost-effectiveness analyses, rather than tradition, anecdotal reports, testimonials, consumer interest, market demand or competition, or political pressures from various interest groups. Answering fundamental questions about CAM therapies requires critical and objective assessment using accepted standards for scientific investigation and rigorous evaluation of scientific evidence. Peer reviewed medical journals provide one

important mechanism for such rigorous evaluation of CAM research, and for providing physicians with reliable information on these diverse and complex therapies.

Dr. Gordon, and members of the Commission, thank you for the opportunity to present these perspectives on publication of CAM research in peer reviewed medical journals. I hope this information is useful in your deliberations on complementary and alternative medicine research and policy.

Clinton Presidential Records
Digital Records Marker

This is not a presidential record. This is used as an administrative marker by the William J. Clinton Presidential Library Staff.

This marker identifies the place of a tabbed divider. Given our digitization capabilities, we are sometimes unable to adequately scan such dividers. The title from the original document is indicated below.

Riley

Divider Title: _____

DAVID RILEY, MD
Integrative Medicine

Profile

Dr. Riley is the Editor-in-Chief of Alternative Therapies in Health and Medicine, a peer reviewed medical journal indexed in the National Library of Medicine. He is the founder of the Integrative Medicine Institute which is conducting practice based outcomes research on effectiveness, safety, patient satisfaction, and costs when complementary and alternative therapies are integrated with conventional medicine. Most recently he was the Chief Medical Office of Medigy, Inc. a company formed to provide on-site and Internet based education in integrative medicine to professionals and consumers. Dr. Riley is currently involved in a variety of clinical research projects from basic science to clinical trials. He co-founded the Integrative Medicine Education Associates (IMEA) a group of physician educators offering conventional physicians and other licensed health care providers an introduction to Integrative Medicine with CME accreditation from Columbia University and the University of Arizona.

Dr. Riley's primary interest in medicine is improving patient care through the evidence-based integration of conventional and alternative therapies. He attended Georgetown and the University of Utah Medical Schools and graduated from the University of Utah School of Medicine in 1983. His undergraduate studies were at the University of North Carolina, where he graduated in 1976 with a BA in music and honors in psychology. He is currently a Clinical Associate Professor at the University of New Mexico Medical School, board certified in Internal Medicine and a certified yoga instructor. In 1993 Dr. Riley completed a three-year training program in homeopathy at the Hahnemann College of Homeopathy in Albany, California..

He has conducted more than 65 clinical trials. He was the principle investigator on the first International Integrative Primary Care Outcome Study. This multi-center outcomes study evaluated patient outcomes when complementary and alternative therapies were integrated in primary care. He has been involved in the development and implementation of data collection networks collecting practitioner and patient outcome and satisfaction data where complementary and alternative medical therapies are integrated into patient care. Southwest Health Options, co-founded by Dr. Riley, was an independent practice association (IPA) managing the delivery of complementary and alternative medicine (CAM) for insured patients in Santa Fe, New Mexico. As the medical director of Southwest Health Options, Dr. Riley directed the peer review and utilization management, and developed credentialing standards acceptable to both the insurance industry and the CAM community.

Dr. Riley is a board member of the Homeopathic Pharmacopoeia of the United States (HPCUS), a technical advisory board to the FDA for regulatory issues concerning homeopathy. Dr. Riley lectures and consults internationally on a range of healthcare issues including education, research, and regulatory affairs.

P.O. Box 4310 o Santa Fe, NM 87502 o USA
Telephone: (505) 983-0546 • Fax: (505) 983-4196
E-mail:dsriley@ibm.net• Internet site: http://www.integrativemed.org

DAVID RILEY, MD
Medicine

Integrative

PROFESSIONAL EXPERIENCE

MEDICINE

- Editor-in-Chief: *Alternative Therapies in Health and Med.* (1999 – 2000, 2001 - present)
- Medical editor: *Alternative Therapies in Health and Medicine* (1995 - 1999)
- Chief Medical Officer, MED*igy*, Inc. (November 1999 – July 2000)
- Private practice: general internal medicine, complementary medicine (1986 - 2000)
- Founder, Southwest Health Options IPA (1997 – March 2000)
- Clinical Associate Professor, UNM medical school (1987 – 2000)

MEDICAL RESEARCH

- Medica/Research Director, Integrative Medicine Institute (1992 – present)
- Gene expression testing of pharmaceutically based CAM products (2000 – present)
- CAM data collection network (2000 - present)
- randomized controlled clinical trial on the treatment of herpes (1999)
- controlled clinical trial on smoking cessation (1998)
- Integrative Med. Data Collection Network, (1998 – present)
- International Integrative Primary Care Outcome Study (IIPCOS-1: 1995 - 1998)
- clinical trial in homeopathy 60 homeopathic drug provings (1992 - present)
- controlled clinical trial on the treatment of arthritis (1991-92)

TEACHING

Internal Medicine

- Co-founder: *Integrative Medicine Education Associates* (July 2000 - present)
- Clinical associate professor – University of NM Medical School (1987 – present)
- Clinical associate professor – University of Utah (1986 – 1989)
- UNM medical students & residents rotate through practice (1987 - 2000)
- St. Vincent's Hospital staff, Santa Fe, NM (1987 – 1993)

Integrative Medicine

- hatha yoga intensives/therapeutic yoga: Pune, India (July,1985 and July, 1987)
- public yoga classes and/or private therapeutic yoga instruction (1981 - present)
- international lecturer on clinical research (1990 - present)
- homeopathy in primary care, a category I CME course for physicians (1996 - present)
- medical residents and students rotate through practice (1987 - present)

Counseling

- trained in conflict resolution, communication skills, and counseling (1983 - 1990)
- facilitated groups & training programs (1989 - 1990)
- organized and facilitated personal growth workshops (1984 - 1992)

P.O. Box 4310 o Santa Fe, NM 87502 o USA
Telephone: (505) 983-0546 • Fax: (505) 983-4196
E-mail: dsriley@integrativemed.org o Internet site: http://www.integrativemed.org

#44

MEMBERSHIPS

- diplomate, American Board of Internal Medicine (1987 - present)
- HPCUS board member: technical advisory committee to FDA (1992 - present)
- member: AMA, NM Medical Society, AIH
- medical licensure: New Mexico (1986 - present), Arizona (1995 - 2000)

OTHER

- professional musician and chef in Europe and US; sculpture (bronze casting, welding, stone), bio-dynamic gardening

EMPLOYMENT

1989 – present	private practice, clinical research, consulting
1999 – 2000, 2001	editor-in-chief, *Alternative Therapies in Health and Medicine*
1999 – 2000	chief medical officer, MEDigy, Inc.
1995 - 1999	medical editor, *Alternative Therapies in Health and Medicine*
1986 - 1989	medical director - New Mexico State Penitentiary, Santa Fe, NM supervised 3 clinics, one hospital unit and 55 employees
1983 - 1986	intern and resident - Internal Medicine, Salt Lake City, UT

EDUCATION

Internal Medicine	residency: Internal Medicine (1984-1986), LDS Hospital, SLC, UT internship: LDS Hospital (1983), Salt Lake City, UT
Medical School	University of Utah, Salt Lake City, UT (1980 - 1983),
M.D.	Doctor of Medicine, University of Utah Medical School - 1983 Georgetown University, Washington, DC (1979 - 1980)
Homeopathy	Hahnemann College of Homeopathy, Albany, CA (1991-1993) (numerous conferences and workshops from 1987 - present)
Yoga	certified Iyengar yoga instructor: beginning teaching certification 1985, junior intermediate teaching certification 1987
Pre-medical	University of North Carolina, Chapel Hill, NC (June 1977 - July 1978)
B.A.	Music and Psychology with honors in Psychology , UNC-CH (1976)
High School	Culver Military Academy, Culver, IN, 1970

SELECTED PUBLICATIONS AND PRESENTATIONS

• Member of the CONSORT group. **The CONSORT Statement: Revised Recommendations for Improving the Quality of Reports of Parallel-Group Randomized Trials.** David Moher, MSc; Kenneth F. Schulz, PhD, MBA; Douglas Altman, DSc; for the CONSORT Group. JAMA. 2001;285:1987-1991.

• **Southwest Health Options (SHO): Experience with a Complementary and Alternative Medicine IPA for Patients with Insurance in Santa Fe, New Mexico.** David Riley, MD. Harvard/UCSF research conference in Complementary and Alternative Medicine. San Francisco, CA, May 17 - 19, 2001.

• **Application of Gene Expression Technologies for the Evaluation and Comjparison of CAM and Conventional Pharmaceutical Therapies.** Danute (Bunki) M. Bankaitis-Davis, Ph.D., David S. Riley, M.D., et al. Harvard/UCSF research conference in Complementary and Alternative Medicine. San Francisco, CA, May 17 - 19, 2001.

• **Faculty member: Kaiser Institute - Fellowship in Integrative Medicine**. Jacksonville, Florida, February 13 - 14, 2001.

• **A Randomized Clinical Trial on the Treatment of oral herpes with Topical Zinc oxide/glycine.** B Godfrey, N Godfrey, J Godfrey, DS Riley. Alternative Therapies in Health and Medicine. May 2001, Vol 7(3).

• **Homeopathy and Conventional Medicine: An Outcomes Study Comparing Effectiveness in a Primary Care Setting**. David Riley et al. Journal of Complementary and Alternative Medicine Research. March, 2001 Vol 7(2).

• Anthroposophical medicine research conference - speaker and moderator, Nov. 2000.

• Consort Guidelines Planning Meeting. San Antonio, Texas, April 2000.

• Medical Challenges for the Millenium - Technology, Ethics, Alternatives. Humboldt University International Scientific Symposium, Institute for Social Medicine and Epidemiology, Berlin, Germany, September 1999. Symposium publication.

• Outcomes Research in Complementary and Alternative Medicine. Keynote Speaker: Carstens Foundation Annual Investigator Conference, Ulm, Germany, September 1999.

• Natural Healing Therapies. U.S. Consultant, Dorling Kindersly, New York, 1999.

• Manuel Lipson Distinguished Lecture. Harvard Department of PM&R , Spaulding Rehabilitation Hospital, Boston, Massachusetts, May 10, 1999.

• Integrative Medicine. Presentation to the Lutheran Medical Center Foundation. Denver, CO April 19, 1999.

• *Alternative Therapies in Health and Medicine Conference*, New York, New York March 1999:
  - An Introduction to Homeopathy for Primary Care Provers
  - Hatha Yoga: Therapeutic Yoga and Instruction in the Practice of Yoga

• Integrative Medicine. Keynote address and workshop presentation. Presbyterian Healthcare System, Presbyterian Hospital, Albuquerque, NM, March 19 - 20, 1999.

• About this Issue: Response to New England Journal of Medicine editorial on Alternative Medicine. *Alternative Therapies in Health and Medicine*, November 1998.

• Integrative Medicine in Primary Care. Presentation. SUNY -Stony Brook Medical School, October 24, 1998.

• Dialogues on High Dilutional Medicine - Mechanisms of Action Theory, # 2. Project Coordinator for the Fetzer Institute; Santa Fe, New Mexico August 23 - 26, 1998.

• The Healing Arts and Science: Complementary and Integrative Medicine. Keynote Presentation. SUNY -Stonybrook Medical School, Vancouver, BC, Canada, July 18 - 21, 1998.

• JAMA - Letter to the editor. May 20, 1998.

• About this Issue. *Alternative Therapies in Health and Medicine*, May 1998.

• Phase I clinical trials - Homeopathic Drug Provings. LIGA conference, Amsterdam, May 1998.

• Symposium Presentations: *Alternative Therapies in Health and Medicine*, San Diego, April 1998.
  - Homeopathy in Primary Care
  - Hatha Yoga: A Practical Introduction

• Encyclopedia of Healing Therapies. U.S. Consultant, Dorling Kindersly, New York, 1997.

• Homeopathy - Implication for Veterinary and Human Medicine. Commonweal conference,

Bolinas, California, November 20 - 23, 1997.

- Integrative Medicine. Keynote address and workshop presentation. Presbyterian Healthcare System, Presbyterian Hospital, Albuquerque, NM, November 14 - 15, 1997.
- Integrating Complementary and Atlernative Medicine into the Health Care Delivery System. Medical staff presentation. Terre Haute Indiana Regional Hospital, October, 1996.
- Introduction to Homeopathy for Physicians. CME program. University ofArizona Medical School, Tucson,Arizona, August 16, 1997
- Workshop: Integrative Medicine for International Physician Educators. Germany, July 1997.
- Dialogues on High Dilutional Medicine - Mechanisms of Action for Homeopathy #1. Project Coordinator for the Fetzer Institute; Kalamazoo, Michigan; July 11 - 14, 1997.
- International, multi-center phase I clinical trial overivew: Proving the Provings:. LIGA conference, Seattle, Washington, May 29 - June 1, 1997.
- Symposium Presentations: Therapeutic Yoga: Implications for manual medicine. Creating Integrated Health Care, *Alternative Therapies in Health and Medicine*, Orlando, FL, April 1997.
- America's Health Network: television interview. Orlando, Florida, April 1997.
- The Mystery of Health: Reclaiming Medicine's Soul. *Alternative Therapies in Health and Medicine*. March 1997. Vol. 3, No. 2, p. 127 - 128.
- CNN: Alternative therapies television interview, November 1996.
- Symptom Selection Criteria - Clinical Verification for Homeopathic Drug Provings. LIGA conference, Capri, Italy, October 1 - 5, 1996.
- Integrative Medicine. Keynote presentation and workshop. Antelope Valley Regional Hospital, Lancaster, California, September 14, 1996.
- Integrative Medicine. Keynote presentation. Columbia/HCA Hospital Association, Colorado Springs, Colorado; August 9, 1996.
- Alternative and Complementary Medicine: What are Our Patients Doing? David Riley, MD and Larry Dossey, MD; Colorado Academy of Family Practice; Aug. 2, 1996.
- Monophasic study design: International Integrative Primary Care Outcome Study. Keynote presentation. International Research Conference, Rosenfeld, Germany; June 15 - 16, 1996.
- Integrative Medicine in Primary Care. Keynote presentation: New Mexico Health Resources annual conference, Taos, NM; June 1996.
- Homeopathy for primary care providers. Seminar presentation: New Mexico Health Resources annual conference, Taos, NM; June 1996.
- Alternative Medicine: Implications for Clinical Practice. Harvard Medical School, Boston, March 1996.
- Symposium Presentations: Creating Integrated Health Care, Sponsored by *Alternative Therapies in Health and Medicine*, San Diego, January 1996.
    - Beyond Technology: The Physician as Healer
    - Hatha Yoga: A Practical Introduction
    - Integrated Use of Homeopathy in Primary Care
    - Research Methodologies: Assessing the Efficacy of Alternative Therapies
    - Working in an Integrative Practice
- Alternative Paradigms in Medicine and Healing Conference, Sponsored by Arizona Center for Health and Medicine & Mercy Integrated Health, Phoenix, September, 1995.
- Original Research - Veronica Officinalis. BHJ, July 1995, Vol. 84, pp. 144-148.
- Original Research - Homeopathic Drug Provings. AANP annual conference, San Diego, September 1994.
- Original Research - Nicotinamide Adenosine Dinucleotide. AIH, Summer 1994.
- Original Research - Myosotis arvensis. BT, July 1994.

- Original Research - Fumaria officinalis. Homeopathic Links; Switzerland, June 1994.
- Thiosinaminum - Three Cases Reports of Infertility and a Review of the potential therapeutic applications of thiosinaminum. IFH case conference, Seattle, WA, June 1994
- Homeopathic Drug Provings. NCH annual conference, Chicago, May 1994. (Reviewed in Homeopathy Today, January 1995.)
- Phase I Clinical Trials and Homeopathic Drug Provings: Rebirth of a Clinical Research Paradigm. AIH 150th anniversary, June 1994
  conference, New York, April 1994; published in the AIH journal, Winter 1994.
- Homeopathy - The Principles and Practice of Treatment. U.S. Consultant, Dorling Kindersly, New York, 1993.
- Introduction to MacRepertory. UCLA continuing medical education , AAMA, Sept. 1991.
- Editorial: Homeopathic Treatment of HIV patients. September 1991, IM World Report.
- Frontiers in Research in High Dilutional Medicine. October 1990, Germany. International conference for researchers in high dilutional medicine. Center for Frontier Sciences, Temple University.
- Conventional and Alternative Treatment of Trauma. Tucson,Arizona; August, 1990.
- Musculo-skeletal Disorders & Biological Therapy, An Integrative Approach. June 1990, Berkeley, CA.
- The Comprehensive Management of Back Pain. October 1987, Chicago, IL.