

Abstract

Context: Various forms of "energy healing" have become popular in the United States.

Objective: To test the assertion that an energy healer can, without physical contact, distinguish the presence or absence of internal organ pathology in individuals who lack overt physical findings.

Design: Observational randomized study, in which we tested the assertion by a well-recognized alternative healer that he had particular skill in using energy transfer to detect the presence or absence of fertility disorders in women.

Patients: Convenience sample of 37 women, 28 of whom had documented pathology resulting in infertility, and 9 of whom were fertile.

Outcomes: The healer was provided with no medical history and performed diagnostic evaluations without physical contact with the blindfolded, clothed, and silent subjects. We compared to random chance the ability of the healer to establish a diagnosis of fertility or fertility disorder.

Setting: Teaching hospital.

Main Results: The healer was unable to distinguish the presence or absence of fertility disorders in the study subjects.

Conclusion: This study points to further need for fair yet rigorous assessment of claims that energy transfer can lead to accurate clinical diagnoses.

Key words: alternative medicine, complementary medicine, energy, psychic, spiritual healer

Spiritual healing encompasses a variety of techniques involving the intentional influence of 1 or more persons upon another living system without utilizing known physical means of intervention.^[1] The force behind spiritual and energy diagnosis and healing is frequently described as "an energy transfer" between the healer and patient, but the idea that an as yet unidentifiable "vital energy" source exists is a major point of contention between proponents of these therapies and clinical investigators. The scientific method argues that if such an energy field exists, it should be measurable by physical means and clearly demonstrable in clinical research studies.^[2]

A nationally representative telephone survey of 2055 adults in 1997 found that 7% reported the use of "spiritual healing by others" (as distinct from personal prayer), and 4% reported using "energy healing" (examples included magnets and laying on of hands) to treat their "most serious or bothersome medical conditions."^[3] Estimates of the number of Americans who sought a spiritual healing or energy healing provider (ie, "healer") for therapeutic sessions in 1997 exceed 2 million individuals per year.^[3]

Prior published investigations suggested that accurate clinical diagnosis using the techniques of iridology^[4] and therapeutic touch^[5] could not be substantiated. We could find no scientific evidence to support claims that "energy diagnosticians" can indeed make accurate clinical diagnoses and designed a study to evaluate the diagnostic accuracy of a recognized healer in active practice. This healer asserted that he could accurately determine the presence or absence of fertility disorders in women who lack overt physical findings.

Methods

We interviewed 5 individuals who, for more than 10 years each, had worked professionally as "energy diagnosticians and healers." They represented an array of healing traditions from China, India, the Middle East and Europe, and each provided care to large numbers of patients. From these 5, we selected a healer who had prior experience working in conventional medical research settings, claimed strongly that he had special skill in the area of infertility, and was willing to participate in an investigation in which he would be blinded to the clinical history and pathologic findings in the study subjects.

An obstetrician/gynecologist specializing in infertility (MS) referred patients with documented fertility problems and patients who could serve as healthy controls. The group consisted of a convenience sample of 28 women with documented anatomic pathology causing infertility and 9 women of similar age with a history of normal conception and delivery of a child 12-36 months prior to the study. All subjects were followed clinically by one of the authors (MS).

We established the following procedures:

- The diagnostician examined experimental and control subjects in a routine examination room in the hospital.
- Subjects wore street clothes (as was this diagnostician's usual practice) and were blindfolded during the entire examination period. This was done to minimize the possibility that the subject could bias the diagnostician's visual cues.
- As was his custom, the diagnostician did not touch the patient, but "scanned" the patient's body using both hands held at a distance of 1-12 inches from the subject's clothes.
- Assessments lasted 3-5 minutes and were videotaped in their entirety.
- During evaluations, there was no verbal communication, with the exception of occasional directions from the diagnostician positioning the patient (eg, "please stand," "please sit," "please move backwards," etc). This, too, was customary for this diagnostician.
- After completing each evaluation, the diagnostician documented his assessment of a patient's medical condition with a series of pictorial representations and a brief, open-ended, written summary of overall findings. In addition, we asked him to answer "yes" or "no" to each of 4 questions for each subject:
 1. Is there an abnormality pertaining to fertility?
 2. Is there a uterine abnormality?
 3. Is there an abnormality involving the pituitary-ovarian axis?
 4. Are there ovarian or fallopian tube abnormalities?

The diagnostician suggested that the experiment focus on fertility issues, as this was a clinical area for which he considered himself to be an expert. With the full consent of the diagnostician, questions 2-4 were added by coauthors in an effort to further explore the healer's diagnostic capability. The healer was not told the percentage of fertile or infertile women, but rather that there would be "some number of each." In an effort to avoid confounding, a physician uninvolved with the study and naive to the study subjects' histories coordinated all patient interactions. The protocol received approval from our hospital's Committee on Clinical Investigations. All patients signed an informed consent form and participated voluntarily.

We computed the sensitivity and specificity of the diagnostician's response to each of the 4 questions. We also computed his accuracy, defined as the sum of sensitivity and specificity, and its 90% confidence interval.^[6] Under the null hypothesis that the healer could do no better than random chance, the expected accuracy is 1, no matter what assumption the healer made about the prevalence of abnormalities in the sample. In order to

assess whether the healer performed better than chance, we calculated a 1-sided P value for accuracy, using the method of Parker and Davis.^[6] The P value in this circumstance derives from the probability that the accuracy is at least as large as was observed, conditional on the diagnostician's assumption regarding the prevalence in the sample of each of the specific conditions. It is calculated based on the binomial distribution.

Results

The Table documents the sensitivities, specificities, and overall accuracy of the diagnostician for each of the questions asked during the evaluation of fertile and nonfertile subjects. The diagnostician's accuracy (ie, number correct/subjects evaluated) ranged from 41% to 58%. Only for the question about uterine abnormality was his diagnosis at least as good as the average performance based on a random guess. For uterine abnormality, though the sum of sensitivity and specificity was larger than 1, the result fails to reach statistical significance ($P = .14$). We conclude that the evidence does not support the healer's claims of diagnostic ability.

Discussion

Despite his claims to the contrary, in this study the energy healer diagnostician was unable to distinguish fertile from infertile women and failed to identify the presence or absence of internal organ pathology in the same subjects. This observation is offered with the caveat that the procedures used deviated from the diagnostician's usual practice, in that subjects were blindfolded and no history was obtained prior to the examination. As such, there is a degree of artificiality that may limit our conclusions.

We designed this study to be simple and pragmatic. It offers a strategy for evaluating alleged abilities attributed to a spectrum of alternative medicine practitioners. Some of their assertions are provocative and, from a biomedical vantage point at times, incredible (eg, that healers can diagnose internal organ pathology without physical contact). As unacceptable as these claims may be to the conventional medical community, the fact that significant numbers of patients accept them as true and pursue them for diagnostic and therapeutic purposes may justify scientific investigation. Moreover, they are testable, ideally in collaborative studies involving qualified researchers and experienced practitioners of alternative therapy (eg, an "energy healer"). Such a collaborative approach can and should be applied to a variety of alternative medical practices.

This study also demonstrates difficulties in designing studies to evaluate questions of this type. For example, there are 2 components of diagnostic accuracy: discrimination and calibration. The former relates to the ability to distinguish between individuals with and without a condition, while the latter addresses, in this case, the diagnostician's belief about the prevalence of the condition in the sample. With the mix of patients in our study, a diagnostician who blindly classified each patient as infertile due to abnormalities of the ovaries or fallopian tubes would have had remarkably high accuracy, but no discrimination. In our analysis, we sought to eliminate calibration as a factor by conditioning on the proportion the healer diagnosed as having the condition. This allowed us to focus on discrimination.

An alternate and perhaps better way to conduct the experiment would be to tell the diagnostician the proportions of fertile and infertile women included in the study. The order in which the women are seen would then be randomized. This would remove from the evaluation assumptions about the prevalence of the condition. The design is then essentially equivalent to R. A. Fisher's classic study of a woman tasting tea,^[7] for which he developed the statistical test that bears his name.

Another design issue is the identification of a comparison basis, ie, with what will the healer's diagnostic abilities be compared? Using a physician who did not communicate with the patient or perform an examination,

we originally designed our study with this physician serving as a control diagnostician. But we realized subsequently that such an individual is essentially a straw man, adding nothing to the evaluation of the healer. It would not be possible to determine whether any difference observed between the healer and the control diagnostician is due to superior discrimination or better calibration (ie, a better guess about the prevalence of a condition within the sample). The appropriate comparison is with random chance, and the question of interest is whether the healer's results differ significantly from what would be expected.

Are age-old claims regarding the diagnostic powers of healers simply age-old fantasy? Perhaps. One can anticipate the argument that such abilities exist, but are diminished or impaired when moved to a research environment in which controlled conditions are strictly maintained. We should note that our diagnostician did not object to the protocol prior to or immediately after he participated in the study.

This study assessed the accuracy of a single diagnostician focusing on a single organ system for which he claimed special competence. Before we can put long-standing assertions to rest, we need additional studies involving a variety of healers and a variety of pathologic conditions. A recent study evaluating the diagnostic accuracy of 21 practitioners of therapeutic touch is an example of the research necessary to evaluate claims in this provocative area.^[5]

A Chinese proverb says, "Real gold does not fear even the hottest fire." If claims of diagnostic accuracy (or therapeutic efficacy) of alternative medical practitioners are true, they will withstand rigorous, collaborative evaluation by and with the conventional medical establishment. If they are false, this too can be discerned through repeated investigation. What is less predictable, however, is the extent to which repeated empiric observations will alter medical providers' beliefs or the propensity to change them.^[8]

References

1. Benor DJ. Survey of spiritual healing research. *Complementary Medicine Research*. 1990;4:9-33.
2. Aldridge D. Is there evidence for spiritual healing? *Advances*. 1993;9:4-21.
3. Eisenberg DM, Davis RB, Ettner SL, Appel S, Van Rompay M, Kessler RC. Trends in alternative medicine use in the United States, 1990-1997: results of a follow-up national survey. *JAMA*. 1998;280:1569-1575.
4. Knipschild P. Looking for gall bladder disease in the patient's iris. *BMJ*. 1988;297:1578-1581.
5. Rosa L, Rosa E, Sarner L, Barrett S. A close look at therapeutic touch. *JAMA*. 1998;279:1005-1010.
6. Parker RA, Davis RB. Evaluating whether a binary decision rule operates better than chance. *Biom J*. 1999;41:25-31.
7. Fisher RA. *The Design of Experiments*. 9th ed. New York, NY: Hafner; 1971.
8. Knipschild P. Changing belief in iridology after an empirical study. *BMJ*. 1989;299:491-492.