

THE EFFECT OF PAIRWISE COMPARISON FORMAT ON CONSISTENCY:
A RANDOMIZED TRIAL.

James G. Dolan, MD

General Medicine Unit, Department of Medicine
University of Rochester School of Medicine and Dentistry

Abstract

Objective: To determine whether the comparison format used for AHP pairwise comparisons affects the consistency of the judgments made.

Design: Randomized trial.

Setting: Ambulatory medical clinic at a University-affiliated, urban teaching hospital.

Population: Twenty volunteers.

Intervention: The participants were randomly assigned to complete an AHP analysis using either the bar graph comparison method or the verbal comparison scale provided by Expert Choice.

Outcome measures: The consistency ratio of judgments comparing the relative importance of five decision criteria to the goal.

Results: The median consistency ratio for the subjects in the graphic comparison group was substantially better than that of the verbal comparison group: 0.027 versus 0.22, $p = 0.007$.

Conclusions: Patients make more consistent comparisons using the bar graph comparison mode. This result suggests that this comparison method is better than the standard verbal scale and should be used whenever possible.

Key words: Analytic Hierarchy Process, Pairwise comparisons, Consistency ratios, Medical Decision Making

Introduction

Over the past several years my colleagues and I have performed a series of experiments to determine whether the Analytic Hierarchy Process (AHP) can be used to improve the process of medical decision making. (1)(2)(3) In the course of this work a number of questions have arisen regarding the best format to use for a "clinical" AHP analysis, i.e., an analysis performed in a brief period of time by someone unfamiliar with this or any other decision making method. Because the consistency of the pairwise comparisons among alternatives and criteria directly affects the validity of an AHP analysis, one of the most important issues for us has been to determine the most appropriate comparison method to use.

Most descriptions of the AHP describe two comparison methods: a numeric scale ranging from 1 to 9, and a verbal comparison scale. In addition to these, Expert Choice, a well known AHP computer program, offers two graphic methods for making comparisons: one that uses bar graphs and another that uses relative areas within a circle. Initially we assumed that patients would have difficulty using the numeric scale and decided to use the verbal scale. In our early pilot studies, however, we found that patients had a great deal of difficulty making consistent comparisons using the verbal scale and did much better when we started using the graphic format. We therefore hypothesized that the use of the graphic comparison scale would result in more consistent judgments than the more standard verbal comparison scale. The purpose of this study, which was done as an adjunct to a study testing the feasibility of using the AHP to assist in clinical decision making (4), was to test this hypothesis.

Methods

Population: The study population consisted of patients attending the Rochester General Hospital Medical Clinic who volunteered to participate in a study of a new method for helping patients play a more active role in making decisions about their health care. Rochester General Hospital is a 533 bed community hospital in Rochester, New York, affiliated with the University of Rochester School of Medicine and Dentistry.

The decision: The AHP analysis used for the study was based on published data describing the pros and cons of alternative approaches to screening for colon cancer in 50 year old men with a family history of colon cancer. (5) The model used for the analysis is shown in figure 1.

The intervention: After a brief introduction, participants were asked to imagine that they were 50 years old, had a first degree relative with colon cancer, and were making a decision about a colon cancer screening program for the next 25 years. They then performed an AHP analysis of the decision using Expert Choice (6) running on a laptop computer. Patients were randomly assigned to either the standard verbal comparison format or the Expert Choice bar graph comparison format using randomly permuted blocks as described by Fleiss.(7)

Data analysis: The study hypothesis was that participants using the graphic comparison format would have lower consistency ratios (indicating more consistent comparisons) for the comparisons among the decision criteria on the first hierarchy level below the goal than those using the verbal comparison format.¹ This hypothesis was tested by comparing the groups' median consistency ratios for the comparisons among the criteria using Mood's median test (8). Statistical significance was defined as a p value of ≤ 0.05 .

Results

Twenty patients (6 men and 14 women, ranging in age from 18 to 63 years) completed the study, ten using the verbal format and ten the graphic format. There were no significant differences between the two groups in age or level of education.

The consistency indices achieved by the verbal and graphic format groups for the comparisons among the criteria are summarized in figure 2. The median consistency ratio achieved by the ten patients in the graphic format group, 0.027, was significantly better than that of the verbal format group, 0.22, $p = 0.007$.

¹ The results of an AHP analysis include a measure of the consistency of the component pairwise comparisons called the consistency ratio. A perfectly consistent analysis has a consistency ratio of 0; higher values reflect increasing amounts of inconsistency. By convention a consistency ratio of 0.1 or less is considered acceptable.

Discussion

These findings indicate that patients using the bar graph comparison format make markedly more consistent judgments among a set of decision criteria than similar patients using the standard verbal comparison scale.

The major limitation of this study is the relatively small sample size. However, as indicated by the high degree of statistical significance, the differences in consistency between the members of the two groups were large and quite uniform. Therefore, I believe these results are much less likely to be due to unique characteristics of the study group than true differences in the performance of the two pairwise comparison methods.

The bar graph comparison format was chosen for this study based on evidence that people are better able to judge differences when they are represented as parallel lines or bars than when they are represented as angles and arcs of a circle. (9) Therefore, the bar graph format should, at least theoretically, be more accurate than the circle graph format which uses the relative areas within a circle to represent the pairwise comparisons. The actual performance of the circular comparison format relative to other formats is, to my knowledge, unknown.

The concept of measurement is fundamental to the AHP. Its great value as a decision making aid is largely due to its ability to enable a decision maker to measure the relative importance of competing alternatives and criteria and subsequently use these measurements to choose among (or prioritize) a set of alternatives. To be truly useful, however, the results of an AHP analysis must be valid. Consequently every effort should be made to ensure that the pairwise comparisons in an AHP analysis are as consistent as possible. The strikingly better consistency found using the bar graph comparison format in this study strongly suggests that this method of making AHP pairwise comparisons is superior to the more standard verbal format and should be used whenever possible.

References

1. Dolan JG, Isselhardt BJ, Cappuccio JD. The analytic hierarchy process in medical decision making: A tutorial. *Medical Decision Making* 1989;9:40-50.
2. Dolan JG, Bordley DR. Involving patients in complex decisions about their care: An approach using the Analytic Hierarchy Process. *J Gen Intern Med*: 1993;8:204-9.
3. Dolan JG, Bordley DR, Miller H. Diagnostic strategies in the management of acute upper gastrointestinal bleeding: Patient and physician perspectives. *J Gen Intern Med* 1993; 8:525-529.
4. Are Patients Capable and Willing to Use the Analytic Hierarchy Process to Help Make Clinical Decisions? Manuscript submitted for publication.
5. Eddy DM. Comparing benefits and harms: The balance sheet. *JAMA* 1990:2493-2495.
6. Expert Choice version 8.0. Expert Choice, Inc. Pittsburgh PA.
7. Fleiss JL. The design and analysis of clinical experiments. New York; John Wiley & Sons; 1986:49-50.

8. Minitab Reference Manual Release 9. Minitab Inc. 1993:18-11.

9. Cleveland WS. The elements of graphing data. Monterey, CA. Wadsworth Advanced Books and Software; 1985.

Figure legends:

Figure 1. The AHP model used for the analysis.

Figure 2. Boxplots showing the verbal and graphic groups median consistency indices. The upper and lower borders of the box represent the 25th and 75 quartile, the line within the box the median. Asterisks are outliers. CI = consistency index.

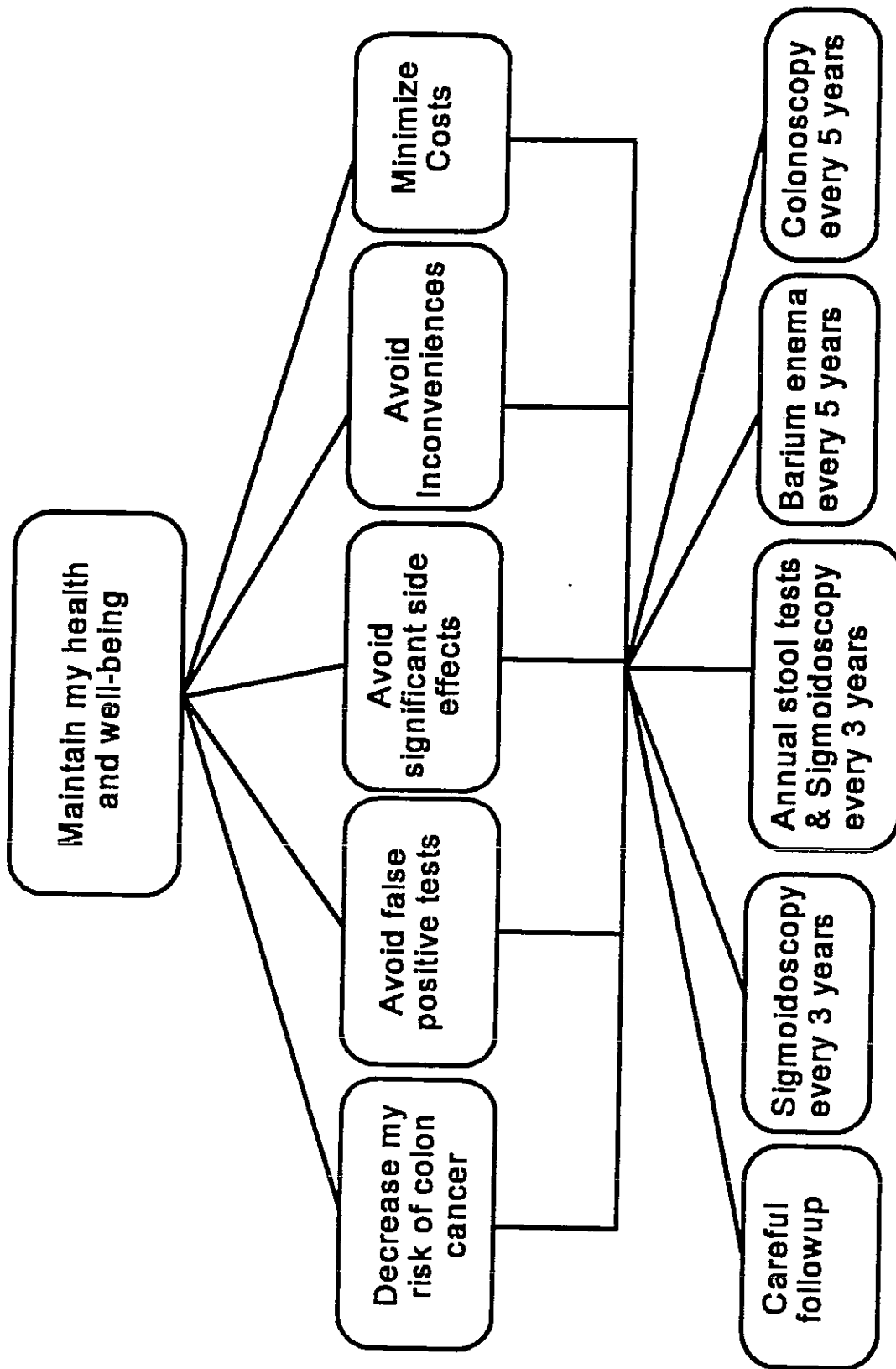


Figure 1

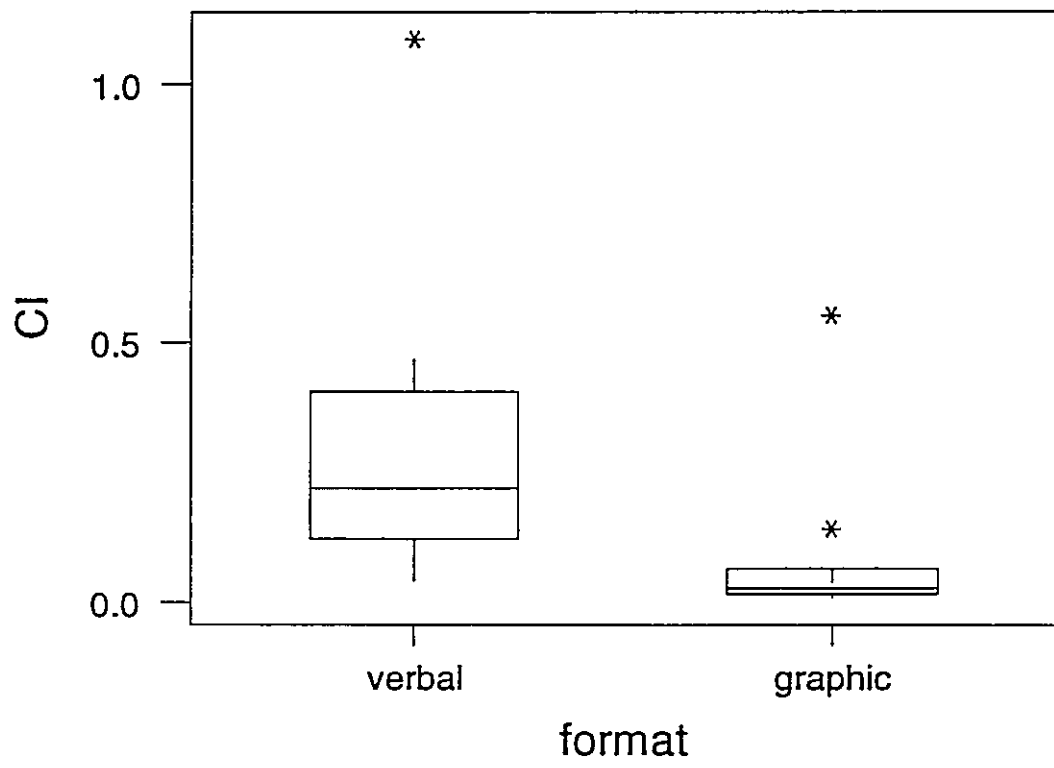


Figure 2.