SEMINAR ON CLINICAL METHODOLOGY

Critical Appraisal of Scientific Literature

Department of Psychiatry, Vanderbilt University

Nashville, Tennessee, USA

June 2, 1994

BASED ON:

Bigelow N, Sainz A. Pitfalls in psychiatric research. Amer J Psychiatry 1962; 118: 889-96.

Chassan JB. Research Design in Clinical Psychology and Psychiatry. New York: Appleton-Century; 1967.

Fisher RA. The Design of Experiments. 6th edition. Edinburgh: Oliver & Boyd: 1949.

Fisher RA. Statistical Methods for Research Workers. Edinburgh: Oliver& Boyd; 1950.

Huntsman AG. Scientific research versus the theory of probabilities. Science 1949; 110: 566.

Hamilton M. Lectures on the Methodology of Clinical Research. First Edition. Edinburgh: Churchill Livingstone; 1961.

Hamilton M. Lectures on the Methodology of Clinical Research. Second Edition. London: Churchill Livingstone; 1974.

Lehmann HE. The placebo response and double-blind study. In: PH Hoch and J. Zubin, eds. The Evaluation of Psychiatric Treatment. New York: Grune & Stratton; 1964.

Ovearall JA, Hollister LE. Psychiatric drug research, sample size requirements for one versus two raters. Arch Gen Psychiatrty 1967; 16: 152.

Wilson EB. Values of statistical studies in cancer patints. Amer J Cancer 1932; 6: 1230-7.

Table 1

Inherent in any research activity is the aim of discovering unknown data; asking a new question and designing an investigation to obtain an answer. Accordingly there are three stages of research:

1. DISCOVERY -- Formulation of hypothesis
   Identifications of appropriate method
   and
   Development of suitable procedures for
   the testing of the hypothesis
   Drawing of inferences from the data obtained

2. VERIFICATION

3. COMMUNICATION

Bigelow, N. and Sainz, A: Pitfalls in psychiatric research. Amer J. Psychiat. 118: 889-896, 1962.

Table 2

Discovery			
Verification			
Communication	Total Information	**1. Review of Literature**	
		2. Statement of Problem	
		3. Method of Investigation	
		4. Experimental Results	
	Dynamic Interaction	**5. Discussion of Results**	
	With Other	**6. Summary**	*Discovery*
	Researcher's Results		*Verification*
			Results

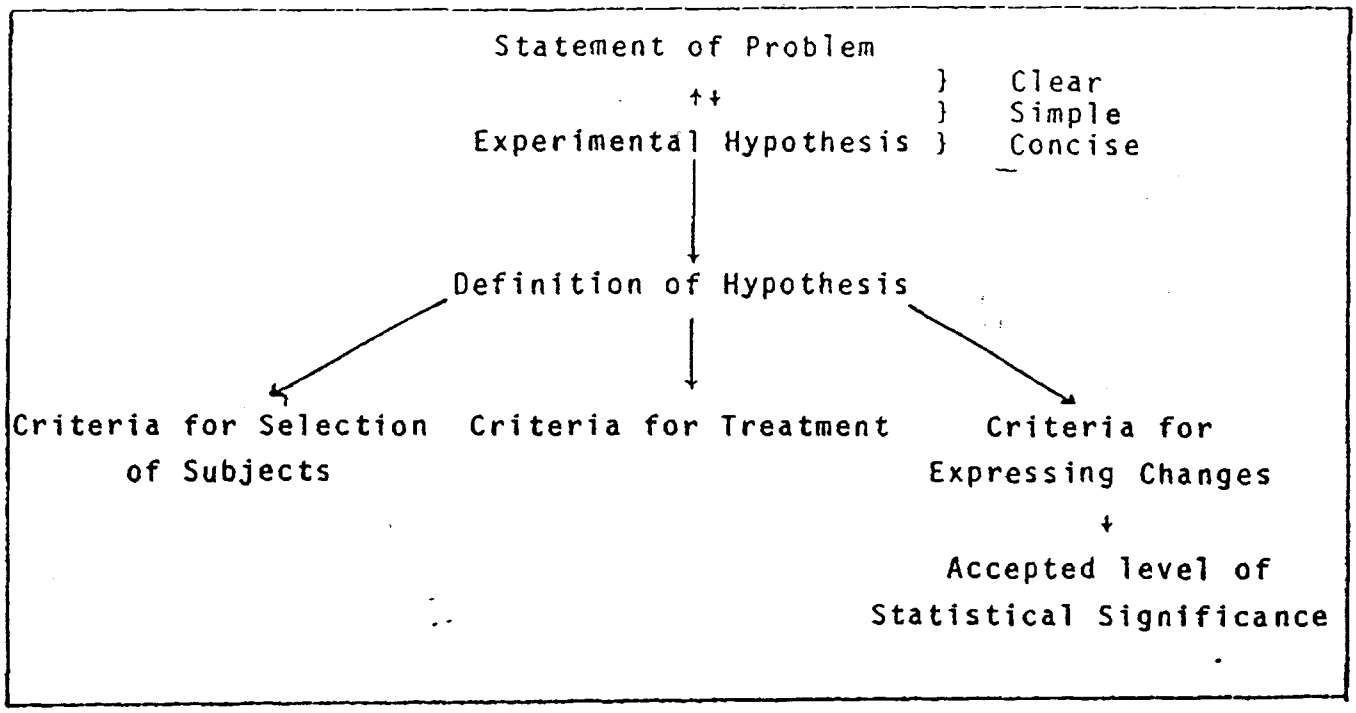
The six sections of an appropriate report

Table 3

Review of Literature	
Prerequisites of Papers to be Used	Prerequisites of a Good Review
Summary should correspond with results	Constructively critical
Results should correspond with experimental design and procedure	Selectively detailed
	Concrete in content

Prerequisites of a good review and of a paper to be used in the review

Table 4



Appropriate statement of the research problem

". . . if we don't know what we are talking about we can still talk and most likely talk volubly, but there is small chance that we are talking to a definite point." (Wilson E.B.: Values of statistical studies in cancer patients. Amer J. Cancer 16: 1230-1237, 1932).

Table 5

It is impossible to prove the null hypothesis, because even if one does not find any difference between two treatments, there is always a possibility that there is a slight difference which would become demonstrable with increase of sample size (Type II or beta error)

It is impossible to disprove the null hypothesis, because it is always possible that the difference between to two treatments due to chance (Type I or alpha error)

It is possible to design the experiment in such a manner that the null hypothesis becomes as unlikely as possible.

(Overall J E and Hollister LE: Psychiatric drug research, sample size requirements for one versus two raters. Arch. Gen Psychiatry 16: 152, 1967.)

Table 6

If one knows that the probability (P) of the null hypothesis is small, at least as low as 0.05, one can safely reject it

Rejection of the null hypothesis implies that one beliefs that the results are due to differences between treatment and not due to chance.

Level of significance of 0.05 means that out of 100 experiments one may be wrong in rejecting the null hypothesis in five.

(Hamilton M: Lectures on the Methodology of Clinical Research. First and Second Editions. Churchill Livingstone, Edinburgh and London 1961 and 1974)

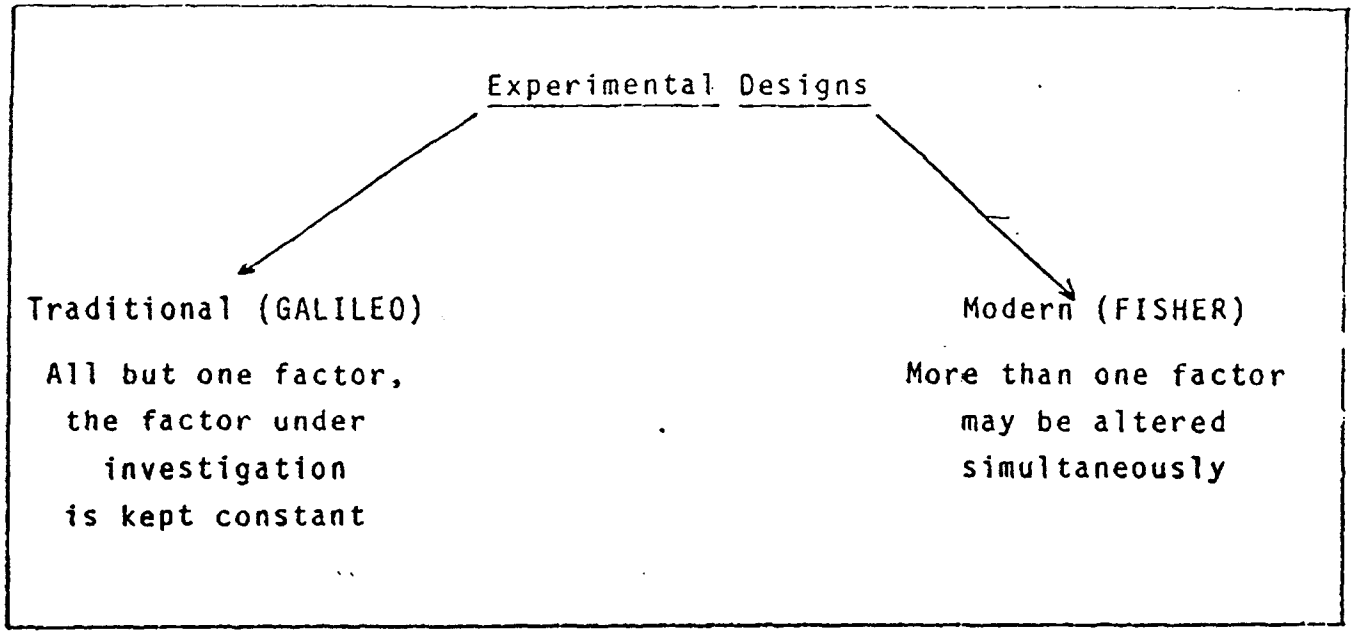
Table 7

To obtain data as quickly as possible using the minimum number of patients for the shortest period of time to obtain the maximum amount of information.

No one wants to withhold a good treatment any longer as necessary bearing in mind the dangers of disseminating unproven or useless treatment.

These goals can be achieved by an appropriate experiemental design relevant to the nature of the question.

Table 8



The essential difference between traditional (Galileo) and modern (Fisher) experimental design

"

(Fisher RA: The Design of Experiments. 6th edition. Oliver and Boyd, Edinburgh 1949; Fisher RA: Statistical Methods for Research Workers. 11th edition. Oliver and Boyd, Edinburgh 1950).

Table 9

PATIENT POPULATION	ASSIGNMENT TO TREATMENT
Selection Criteria	Randomization
Diagnosis	Matching
Prevailing Symptoms	SETTING
Severity of Illness	Private Practice
Duration of Illness	Clinic
Length of Hospitalization	Hospital
Age	ASSESSMENT INSTRUMENTS
Sex	Rating Scales
TREATMENT	Validity
Dosage Form	Sensitivity
Dosage Level	Reliability
Dosage Schedule	
Dosage Regime	
(fixed, fixed/changing, flexible)	
Duration	

Specification of study characteristics: patients, treatments, assignment to treatment, setting(s), and assessment instruments

Table 10

Study Object	Control Requirements
Stable Physical System	→ No Control
Unstable Physical System	→ Needs Control Comparison of the system subjected to the effect of the independent variable with another similar system not exposed to this effect
Complex system that responds to stimuli with varied behavior	→ Needs Control Single-blind Placebo
More complex system that allows for transactional processes	→ Needs Control Double-Blind Placebo

Controls in studying the effect of an independent variable on systems at various levels of complexity

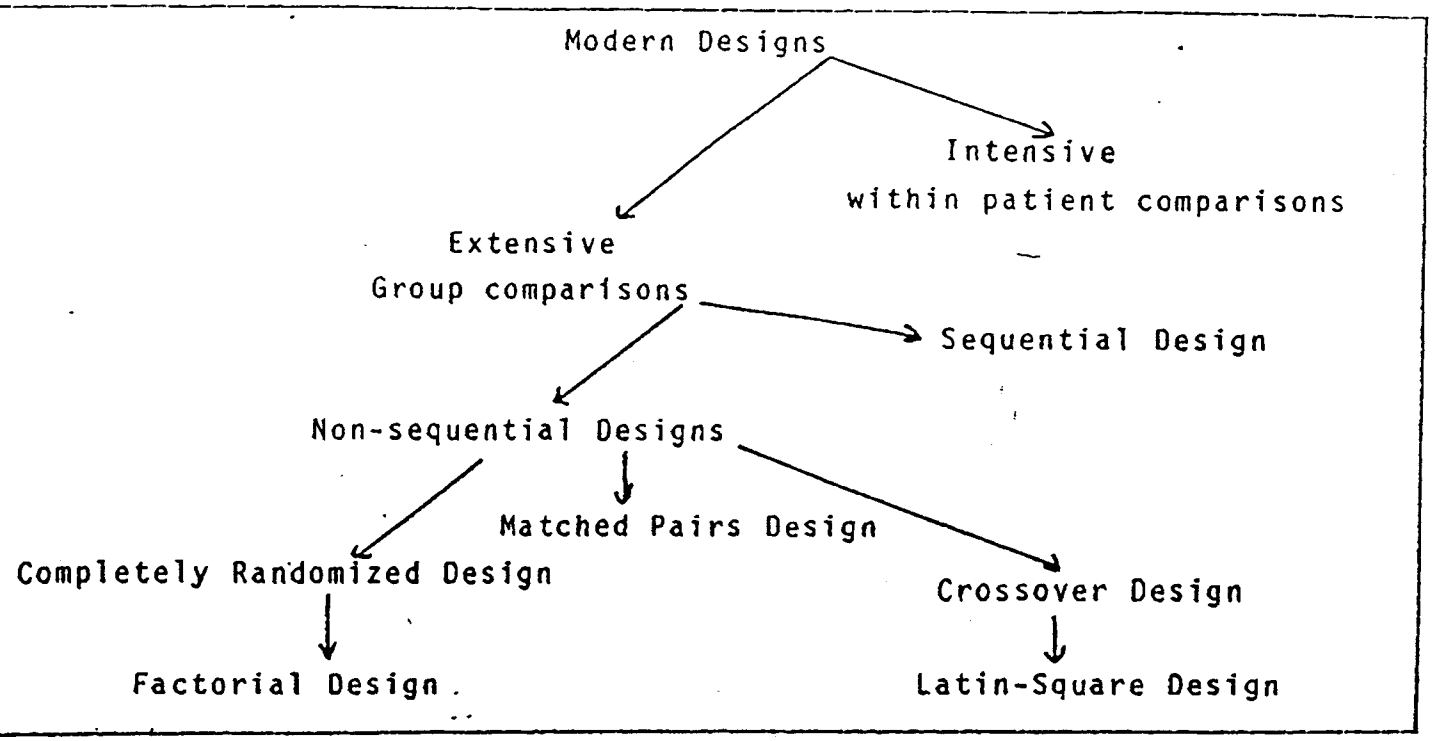
(Lehmann HE: The placebo response and double - blind study. In P. H. Hoch and J. Zubin eds: The Evaluation of Psychiatric Treatment. Grune and Stratton, New York, 1964.)

Table 11

Experimental	
Method(s)	Control(s)
Open- non blind	No drug
Single blind- for patient	Active placebo Inactive placebo
Double blind- for patient for assessor	Standard

Control groups and experimental methods employed in the clinical evaluation of psychotropic drugs.

Table 12



Frequently used experimental designs
in clinical trials with psychotropic drugs

(Chassan J.B.: Research Design in Clinical Psychology and Psychiatry. Appleton - Century - Crofts, New York, 1967.)

Table 13

Treatment O	40
Treatment R	40
Treatment D	40
Total number of cases	120
Comparison of R treatment with no (O) treatment	
Comparison of D treatment with no (O) treatment	

Adopted from Hamilton (1974)

Example of the completely randomized experimental design

Provides information on relative merits of R and D treatments.

Table 14

Treatment 0	30
Treatment R	30
Treatment D	30
Treatment R & D	30
Total number of cases	120

	0	R	Total
0	Tr 0 30	Tr R 30	60
D	Tr D 30	Tr R & D 30	60
Total	60	60	120

Comparison of R treatment with no (0) treatment

Comparison of D treatment with no (0) treatment

Comparison of R treatment with D treatment

Value of interaction between R & D treatments

Adapted from Hamilton (1974)

Example of the factorial design

Table 15

Groups	First Treatment	Second Treatment
Group A	A	B
Group B	B	A

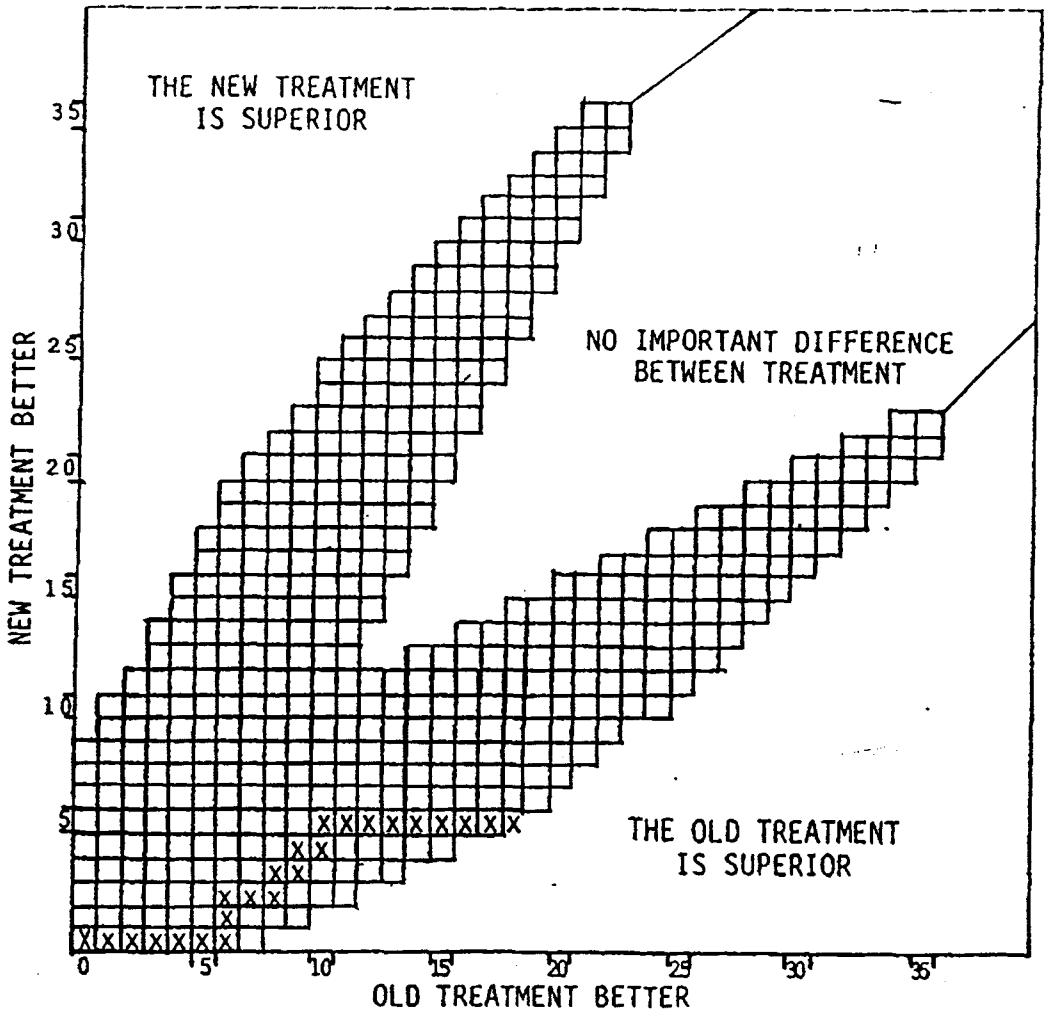
Example of the crossover design

Table 16

	Subject 1	Subject 2	Subject 3
1st Treatment	A	B	C
2nd Treatment	B	C	A
3rd Treatment	C	A	B
Totals	For Subject 1	For Subject 2	For Subject 3

Example of the latin-square design

Table 17



Adapted from Hamilton (1974)

Example of the sequential design

Patients are admitted to the study in pairs (with or without matching) as they present themselves and are assigned to the new or to the old treatment. From each pair the one who responds more favorable to treatment is added to the appropriate line, i.e., old treatment (bottom) line, or new treatment (up left-hand side) line. The trial is terminated when either of the two lines "crosses" into the area New Treatment is Superior or Old Treatment is Superior

Table 18

Appropriately designed controlled experiment gives a valid estimation of experimental error (variation) -- the basis for the statistical tests of significance -- that makes the experiment self - contained

Controlled experiment is necessary because according to the null hypothesis any therapeutic results may have occurred by chance and there is no way of knowing what is the frequacy or rarity of this chance, i.e., the probability that the findings are real.

In the interpretation of data derived from clinical trials inductive logic is used, i.e., members of a class considered and inferences about the class are made on the basis of these members.

The most important kind of induction is the statistical by which general conclusions are drawn on the basis of a limited experimented sample.

Table 19

SIGNIFICANT DIFFERENCE BETWEEN PROPORTIONS (percentages)

Independent Samples	Chi-Square Contingency
Paired Samples	McNemar Test

SIGNIFICANT DIFFERENECE BETWEEN AVERAGES (means) with ASSUMPTIONS that OBSERVATIONS have GAUSSIAN (normal) DISTRIBUTION. Parametric Satististics T-test

Tr vs Random Variation	Analysis of Variance
Tr vs RV vs Initial Scores	Analysis of Covariance

SIGNIFICANT DIFFERENCE BETWEEN AVERAGES (means) with ASSUMPTION that OBSERVATIONS do not have GAUSSIAN (normal) DISTRIBUTION. Nonparametric Satististics

Independent Samples	Median Test Mann-Whitney U-Test
Paired Samples	Sign Test Wilcoxon's Matched - Pairs Signed Ranks Test

Table 20

". . . the prestige of mathematic is so great that many persons forget that even in mathematical hands, probability, chance and random mean ignorance. They come to think that in the alembic of mathematics, chance in some way becomes certainty. They take great care to select random samples without realizing that in so far as a sample has been random, they don't know how it was selected"

(Huntsman AG: Scientific research versus the theory of probabilities. Science 110: 566, 1949.)

Table 21

". . . (clinical investigators) because they are unduly sensitive or insecure regarding their lack of mathematic training and knowledge habitually hand over all their date to biometricians for analysis in order that their papers may include the appropriate chi-square tests, standard error and so on. In that way they have come to depend more and more on mathematicians who have no knowledge or understanding of the subject to intepret their findings, instead of relying on their own experience and common sense" (Wiener, 1962).

Table 22

DISCUSSION

1. To what extent are the statistically significant findings in agreement with ones own clicical judgment and to what extent are the statistically significant findings in variance with the results of others in the literature.

2. Theoretical and practical implications of the findings are considered

3. Recommendations for the clinical implementation of some (or the whole) findings and/or for further research.

IN A GOOD DISCUSSION THE RESULTS BECOME PART OF THE WHOLE EXISTING INFORMATION

Table 23

SUMMARY

1. Brief re-statement of the research problem

2. Short account on the method and procedure employed in studying the problem.

3. Resume of research results

IT IS ESSENTIAL THAT THE SUMMARY IS INFORMATIVE AND THAT IT IS FORMULATED IN SIMPLE TERMS

EDITORIALIZING AND PRESENTING FINDINGS IN THE SUMMARY WHICH ARE NOT FULLY SUPPORTED BY EVIDENCE SHOULD BE CAREFULLY AVOIDED.