DOCUMENT RESUME

ED 136 087 CG 011 094

AUTHOR Harmatz, Jerold
TITLE Tentative Approaches to Testing Elderly Volunteers in a Drug Trial.
SPONS AGENCY Public Health Service (DHEW), Washington, D.C.
PUB DATE 3 Sep 76
GRANT USPHS-MH-12279
NOTE 14p.; Paper presented at the Annual Convention of the American Psychological Association (84th, Washington, D.C., September 3-7, 1976)

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS Drug Therapy; *Geriatrics; *Gerontology; Measurement Techniques; Memory; *Older Adults; *Program Design; *Senior Citizens; *Testing Problems

ABSTRACT

From a study on drug treatment effects on the memory and affective functioning of ambulatory geriatric volunteers, the author recommends several methodological accommodations as necessary and expedient to geriatric assessment. Special procedural attention is required by certain characteristics unique to, or exaggerated among, elderly participants. Designs must acknowledge self-presentation difficulties, symptom overlap, and the importance of keeping things simple. (Author)

TENTATIVE APPROACHES TO TESTING ELDERLY VOLUNTEERS IN A DRUG TRIAL

Jerold Harmatz

Presented at the Annual Meeting of the American Psychological Association, September 3, 1976, Washington, D.C.

From the Psychopharmacology Research Laboratory, Harvard Medical School and Massachusetts Mental Health Center, 74 Fenwood Road, Boston, Massachusetts, 02115. Supported in part by USPHS Grant MH-12279.

ABSTRACT

We have been studying drug treatment effects on the memory and affective functioning of ambulatory geriatric volunteers. From our experience we recommend several methodological accomodations as necessary and expedient to geriatric assessment. Special procedural attention is required by certain characteristics unique to, or exaggerated among, elderly participants. Designs must acknowledge self-presentation difficulties, symptom overlap, and the importance of keeping things simple.

TENTATIVE APPROACHES TO TESTING ELDERLY VOLUNTEERS IN A DRUG TRIAL

I would like to share with you some dilemmas we experienced. We wanted to evaluate psychotropic drugs used to treat early senile dementia. We began with a study of dihydrogenated ergot alkaloids (DEA), viz Hydergine. Clinical practise had repeatedly indicated a great deal of overlap among various kinds of symptomatic experience, including memory loss and the presence of depression, expecially in elderly populations.[1] Before we could turn to the problems of evaluating change, we had somehow to establish pretreatment levels on at least these two dimensions. Otherwise, we needed to have as good an idea as possible about the full spectrum of pretreatment functioning. These requirements of course obtain in treatment assessments of young people, but we encountered what were for us rather novel problems when we turned to study the elderly.

Let me briefly identify certain rubrics under which our problems seemed to fall: perhaps the most formidable difficulty arose from what we labelled problems of self-presentation, what is recognized clinically as denial of symptomatology. The next set of difficulties have to do with the multidimensional nature of both memory and affective dysfunction. A third problem, which complicates the sorting out of these several dimensions, and with which I am sure all of you are all too familiar, is the procedural imperative to keep things simple, to minimize complexity in every experimental task, including the objective monitoring of memory performance-which we judged an important complement of self-reported

decrements. Finally, I want to allude to another obvious fact, which nonetheless easily gets lost in the experimental shuffle, the fact that other things are happening concurrently in the lives of the people serving as subjects for our studies.

A few words about our intended drug study: We used newspaper advertisements to recruit men and women over age 60 who were having problems with their memories (and correspondingly with their affect) but were otherwise healthy. Careful screenings ultimately assured that these criteria were met. We wanted to sort out possible differential responses to the DEA, reported to enhance memory in impaired elderly, in contrast to the antidepressant effects of imipramine and the control of placebo. We randomly assigned individuals, in a double-blind fashion, to participate in one or another of these independent treatment groups. The study was to persist over nine weeks, because something like seven weeks is required hypothetically for DEA to manifest its placebo superiority.

In this first study, we were obliged to depend almost entirely on observer ratings, in particular, the Sandoz Clinical Assessment - Geriatric.[2] I think it is important to recognize that drug intervention can be a most important research tool whereby aging effects can be differentiated by the kinds of co-varying changes seen in various dimensions, e.g., affective state and memory. In this study, both active preparations significantly differed from placebo in their effect on overall clinically rated performance (Slide 1). An accident of sampling resulted in higher depression levels in the initial ratings of people about to receive imipramine. We see here an appropriate covariance adjustment.

It is important for assessment procedures in a given population to focus on symptomatic areas with recognized importance and more than minimal likelihood of occurrence in a population. One does well to avoid the noise introduced by items of mere supposed relevance. For example, in Slide 2 we see what happened to observer ratings in our nine-week trial of the known-effective psychotropic agent, imipramine. Two groups of items are distinguished by how well they differentiate drug-placebo contrast. We can see that the items that do "work" are items reflecting symptoms expectable in the subject population and interpretable by the raters. The non-differentiating items all seem much more desultory and demanding of inference. It is noteworthy that when a discriminate function analysis was applied to the 18 symptom changes experienced by the three drug groups, it was clear that depression and bothersomeness were the most effectively differentiating items.

Now to return to describing our problems: with regard to their denial, seemingly every time we inquired about something "bad," e.g., some dysphoric experience, our respondents would say, "I am not that kind of person." They seemed to take any and all acknowledgements of unpleasant feeling states, of psychological deficit, as a kind of unacceptable typecasting. Ordinary self-assessments like the Minnesota Multiphasic Personality Inventory (MMPI) were next to useless and were universally despised. We have separately examined the MMPI depression item response of elderly participants: we found excellent concordance between our subjects and elderly patients routinely tested at Mayo Clinic, while in both samples, elderly respondents strikingly differed from comparison groups of young people.[3] Presumably, something

like age-cultural expectational sets are operating: it is okay to detail somatic complaints, but these people were brought up, we believe, not to talk about things like sadness, anxiety or anger. Furthermore, these volunteers are terribly competitive, continually making asides about their not being as bad off as others around them, at home and in the study, and always asking if their performance is up to some implicit par.

Two self-rating strategies seem to circumvent much of this kind of thing: one wherein the assessment derives from default, from failure to endorse, e.g., happy items even when sad items are simultaneously passed over; this is exemplified by Depression Adjective Check List.[4] In Slide 3 we can see the expected drug effect, especially of imipramine. The reduced N is a consequence of our having introduced this measure only belatedly into our design. A second device we have recently used, and much prefer for its directness and flexibility, is what is called the line test (Slides 4-7). Subjects self-assess their psychological states by marking how they feel on a 100 mm line, for which the anchor points can be readily defined, but which avoids the semantic encumbrance of Likert-scale modifiers. Each line serves as a visual analog of a given feeling, while there is no need for quibbling over terms like "moderately" or "somewhat."

With regard to memory assessment, we learned that the customary measures of intellectual functioning just would not work to discriminate sicker from healthier responses among these people. Similarly, tests of presumed organicity, e.g., the Goldfarb[5] we felt would not be suitable for these relatively intact participants.

We have had also to develop our own indices and tests of some of the different processes subserving memory, or if you will, different aspects of memory function. For example, we saw that very short term memory, as well as the accessibility of more crystalized material, differed according to the personal meaningfulness of the content. When we had used the Wechsler Form I logical memory story of the lady from South Boston who was robbed, etc., our subjects could remember the story much better than they could the second story about a ship striking a mine (Slide 8). We have recently begun to capitalize on this differential recall of auto- and allocentric material: for example, we have developed the FASOT (Slide 9). I'll briefly describe its administration. Among persons with impaired memory, our hypothesized recall-superiority of personal information seems supported.

Related both to the issues of self-presentation and the complexity of what is being investigated, is the problem of keeping the given experimental task we oblige a subject to perform as simple, as transparent, as straightforward as possible. We have been trying, for instance, to measure life stress in our people; the Holmes and Rahe[6] test involves a comparison operation, viz., "compared to the stress of marriage, would you say troubles with your work are more or less stressful." None of our impaired volunteers, despite their everyday ambulatory living and self-care, has been able to make this kind of comparison; they rapidly stimulus-displace, and respond with an absolute indication of how stressful work might be.

Any gadgetry, even a stopwatch, intimidates them, provoking their

self-esteem and prospective health concerns, and serves to distract them from the substantive task. We use an interference nonsense syllable recall test patterned after Broadbent[7] but we avoid the split-screen slides and headphones usually used in such tasks (Slide 10). All of our tests have had to compromise some experimental integrity by accomodating hearing loss in one subject and visual decrement in another.

Then there is the inescapable reality that lots of things happen to these folks over the course of the few months they participate with us. We have had to build into our procedure careful attention to these concurrent events for one, the hospitalization of a spouse, for another, the sale of some jewels and sudden realization of some needed income. Such changes in life circumstance have powerful, but not always obvious, effects on treatment assessment.

In general, we would argue for a quality of naturalness in our evaluation procedures, a degree of flexibility that keeps us from locking ourselves into an experimental (e.g., multifactorial) design which might be too unwieldy or impracticable, or perhaps wide-open to unarticulated artifact. From the vantage of statistical reasoning, it will be useful to employ multiple regression procedures, explicitly declaring the presence or absence of particular qualities or risks, so that they come into the ultimate variance reckoning.

Very recently we have designed a study in which we hope to benefit from innovative applications of time series analysis.[8] We all readily appreciate the terrible difficulties of selecting comparable numbers of patients to participate in a long-term traditional independent group, controlled experimental plan. It is possible that a great deal can be learned from an intensive study of particular individuals. Proper

inferences, however, are possible only if the resulting data are subjected to suitable statistical treatment to identify and isolate autocorrelation effects. We have begun to train individual elderly volunteers to rate themselves by telephone on some specially modified visual analog scales. Each subject will make alternate-day observations for two months before and two months after active drug treatment intervention. Both intervention and covariance effects, e.g., differential effects in memory and depression, we hope will be detectable.

I frankly doubt that our experience is all that unique. Perhaps other participants will be able to relate to what I have depicted as problems and strategies attending a psychotropic drug trial.

STUDIES WITH ELDERLY SUBJECTS

COVARIANCE ADJUSTMENT OF TOTAL TREATMENT CHANGES

SCAG TOTAL

	Hyd	Imip	Pbo	F-Test	P
$\bar{X}$	-11.500	-14.846	-4.286	6.144	.006
SD	7.106	6.926	8.398		
$\bar{X}'$	-11.680	-15.105	-3.847	4.924	.014
N	10	13	14		

STUDIES WITH ELDERLY SUBJECTS

OBSERVER RATINGS (SCAG) SENSITIVE TO DRUG EFFECT

(IMIPRAMINE-PLACEBO CONTRAST)

Discriminating Items	univariate t-test	Non-discriminating Items	univariate t-test
Confusion	1.98	Alertness	0.75
Anxiety	2.08	Memory	0.42
Depression	3.10	Disorientation	1.76
Irritability	2.14	Lability	0.37
Hostility	2.08	Initiative	0.86
Bothersomeness	2.76	Indifference	1.00
Unsociability	2.49	Self-Care	1.00
Uncooperativeness	2.07	Fatigue	0.81
		Anorexia	1.41
		Dizziness	0.69

Hotelling T^2 = 49.8094

P = .004

Hotelling T^2 = 15.8519

P = .472

STUDIES WITH ELDERLY SUBJECTS

NINE-WEEK EVALUATION OF HYDERGINE, IMIPRAMINE AND PLACEBO CHANGES

Depression Adjective

Check List (DACL)	Hyd	Imp	Pbo	F-Test	P
$\bar{X}$	-2.000	-4.333	1.818	3.708	.038
SD	3.391	7.192	3.242		
N	8	12	11		

References

1. Kahn, RL, Zarit, SH, Hilbert, NM, Niederehe, G: Memory Complaint and Impairment in the Aged, Arch Gen Psychiat 32:1569-1573, 1975.

2. Shader, RI, Harmatz, JS, Salzman, C: A New Scale for Clinical Assessment in Geriatric Populations: Sandoz Clinical Assessment-Geriatric (SCAG), J Am Geriat Soc 22:107-113, 1974.

3. Harmatz, JS, Shader, RI: Psychopharmacological Investigations in Elderly Volunteers. II. MMPI Depression Scale, J Am Geriat Soc 23: 350-354, 1975.

4. Lubin, B: "Manual for the Depression Adjective Check Lists," San Diego, Education and Industrial Testing Service, 1967.

5. Kahn, RL, Goldfarb, AI, Pollack, M, Peck, A: Brief Objective Measures for the Determination of Mental Status in the Aged, Am J Psychiat 117:326-328, 1960.

6. Holmes, TH, Rahe, RH: Social Readjustment Rating Scale, J Psychosom Res 11:213-218, 1967.

7. Broadbent, DE, Heron, A: Effects of a Subsidiary Task on Performance Involving Immediate Memory by Younger and Older Men, Brit J Psychol 53: 189-198, 1962.

8. Glass, GV, Willson, VL, Gottman JM: Design and Analysis of Time-Series Experiments, Boulder: Colorado Associated University Press, 1975.