UNIVERSITY OF CINCINNATI
DEPARTMENT OF PREVENTIVE MEDICINE AND INDUSTRIAL HEALTH
COLLEGE OF MEDICINE
CINCINNATI 19, OHIO

August 25, 1959

MEMORANDUM:

TO: Dr. Robert A. Kehoe

FROM: Biometrics

SUBJECT: Design of Lead Study

Purpose

1. The purpose of the experiment will be to obtain reasonable estimates of the various effects that would result from subjecting humans to lead-in-air for prolonged periods of time.

2. The research design should enable the investigator to predict:

   a. An estimated value of lead in urine, feces, blood, tissue, etc. after the subject breathes different concentrations of lead or breathes them for prolonged periods of time, and;

   b. The error with which such predictions can be made.

3. It is not possible to fulfill the aims of the experiment without commitment to a long term study involving a number of subjects under specified conditions of testing. Unfortunately the analyses and designs of long term experiments are not yet in advanced stages. Some such experiments have been performed in agronomy with mixed successes. The specific design suggested here will incorporate design features that have proven successful in the past. However, the problem of estimating effects to projected experimental conditions demands some new approaches. The solution advocated here is to use à priori predictions as guides to generating estimating functions. The à priori prediction will consist of reasonable hypotheses concerning the general shapes of experimental effects. Only part of the data will be used to choose the best prediction equation which will be based equally on the hypothesized form of that equation and the best fit to observed data. The rest of the observations will be used to select among alternative hypotheses and to derive estimates of prediction errors.

Design

1. The design of the experiment should consist of a counterbalanced series of subject exposures to different concentrations of lead in air, or, for different periods of time to fixed concentrations, or, to a series of exposures in which both concentration and duration are varied. The choice of fixing either concentration or duration may be determined by experimental practicalities. However, it should be pointed out that if both duration and concentration can be varied concurrently, the ultimate number of subjects needed will be smaller than if each variable is varied in separation, interactions between duration and concentration can be evaluated without the usual penumbra of uncertainty, and estimation of prediction error will be improved.

2. After discussing practical procedures and feasible experimental conditions with Mr. Schafer, from the evaluation of data collected in previous experiments, and with the expectation of utilizing that data as part of the final analysis it appears that the following combinations of conditions and treatments will make for the most efficient design:

   a. Each subject can be used as a complete experiment. Optimum efficiency is obtained if the final number of subjects is some multiple of 4. (i.e. 4, 8, 16, etc.)

   b. Two sets of experimental conditions should be varied concomitantly:

      i. Concentration of lead in air: Values to be used to vary from 75 γ/m^3 to 125 γ/m^3, 150γ/m^3, and 175γ/m^3. Mr. Schafer estimates the time needed to modify the equipment (to permit daily variations of concentrations) not to exceed one month.

      ii. Durations of experimental sessions: Length of sessions to be divided into two groups:

          Short - term sessions: 2, 4, 6, and 8 hours,
          Long - term sessions: 9, 10, 11, and 12 hours.

      The ratio of short to long term exposures could be 3:1. The subject would average slightly less than 7 daily exposure hours. Overtime, ranging from one to four hours would occur either once or twice each week. (If this schedule should prove too onerous for the subject the ratio of short to long term sessions could be changed to 4:1 without appreciably changing the total design.)

c. Counterbalancing should be such that serial effects of experimental sessions can be separated from specific treatment effects. The most efficient design for that is a variation of orthogonal Graeco-Latin Squares. The specific block design will be such that each treatment level precedes and follows all other treatment levels. Successive blocks may be treated as within subject replications and advantage can be taken of the convenient analogy between Latin Square design and confounding in factorial experiments.

d. The number of experimental sessions is given by the minimum number of long term exposure sessions needed for maximum efficiency. This number is 32 and since the ratio of short to long term exposures is 3:1, the total experiment should consist of 128 experimental days. (For a 4:1 ratio of short to long term sessions, the total number of experimental days would be 160.)

e. The first block of 16 short term and 4 long term sessions in the planned quasi orthogonal Graeco-Latin Square are given below

Week	Duration in Hours	γ/m^3
	2	75
	4	125
1	6	150
	12	75
	8	175
	4	175
	2	150
2	11	125
	8	125
	6	75

Week	Duration in Hours	γ/m^3
	6	125
	10	150
3	8	75
	2	175
	4	150
	9	175
	8	150
4	6	175
	4	75
	11	175
	2	125
	etc.	
5		

f. Procedure: Air intake will be prescaled so that a desired concentration can be "dialed" for each experimental day. The room will be made ready at 8 A.M. and the subject will start each day at 9 A.M. All experimental sessions will be continuous except for a 10 to 20 minute break each four hours to permit the subject time to eat and to refresh himself. (All nourishment is to be taken outside the room after a thorough mouth-washing.) Urine, feces, blood, and samples of food will be collected as was done previously with one exception. The subject will be asked to

empty his bladder at the end of some experimental sessions and collect his urine for the next 48 hours in individual containers - on each of which he is to mark the time of urination. This procedure is to be followed for at least 8 to 12 weekends.

Estimation

1. Component variables will be evaluated for effects and possible interactions. Because of the particular counterbalancing of the Graeco-Latin Square design this can be done very efficiently. After the first part of the analysis selects the variables that appear to affect, either singly or in combination, the values of lead excreted in urine or feces or the amounts found in the blood, these variables will then be used as estimators or predictors for the second part of the analysis.

2. Estimation of effects to projected experimental conditions will be based on the best fitting function that can be derived from the data under the following specific conditions:

   a. No more than 2/3 of the data will be used to derive a prediction equation;
   b. The general form of the prediction equation will be specified by an à priori hypothesis;
   c. Fitting the equation itself will be limited to points derived from short time experimental exposures.

   No more than 2/3 of the data will be used in deriving equations so that the remaining 1/3 can serve as check on the validity of the final function.

   The specific and general form of the prediction equation should be determined more by physiological considerations than curve fitting. In this way the final prediction will not only suit the collected data but will also carry the conviction of being physiologically reasonable. Curve fitting procedures will be limited, therefore, to the determinations of constants for an otherwise independent quantitative hypothesis.

   The interaction between fitting the final curve and use of a priori hypotheses needs further elaboration. The à priori statement should describe a reasonable physiological expectation. For example, one could postulate that excretion of lead in urine approaches some physiological limit which remains fixed, or that excretion of lead in urine approaches a physiological limit that is itself determined by other conditions, or that excretion of lead in urine approaches a physiological limit that increases constantly with adaptation to higher levels, and so on. Each of these statements can be described in general quantitative

forms. A simple quantitative translation of the statement that the physiological limit of lead excretion remains fixed would be

$E(lu) = U(1 - e^{-ad}) \pm p_e$; where $E(lu)$ = expected urine lead;

U = the physiological limit, A = a physiological constant,

d = duration of exposure, and p_e = error of prediction.

Similar quantitative statements can be made to express the other alternatives. Within these quantitative expressions are still other alternatives and the process of analysis will consist of a stepwise elimination of non-predictive forms of equations. With present computer facilities such a project does not appear to be excessively difficult.

Fitting the first approximation equations only to the effects of short duration experiments is the first step in the elimination of poor approximation. Since the problem essentially consists of deriving estimates of the effects of long duration exposures from observed shorter duration exposures, the method of analysis will attempt to predict the observed effects of longer durations from the observed effects of shorter durations. In this fashion it will be possible to select among the possible quantitative statements those that, when fitted to short duration experiments, will also predict long duration experiments.

If the different experimental exposures range from 2 to 12 hours then only the data for durations of no more than 6 hours will be used to determine the constants in the equation. Alternative equations will be rejected or selected according to the size of the error with which they predict the effects of experimental durations of longer than 6 hours.

3. Two safeguards are introduced by this method of analysis. First the form of the equation that gives the best fit is originally determined by physiological considerations. Thus the expectations are, from the beginning, that with the correctly phrased equation one may expect to fit all possible points on the curve. Secondly, the data ~~are~~ used in part as if effects of longer duration exposures, (from 8 to 12 hours) although observed during the experiment, are to be predicted - as well as the effects to exposure durations which have not been observed (e.g. from 13 to 24 hours). It is reasonable to assume that if the method yields satisfactory estimates of observed phenomena (i.e. for 8 to 12 hours) it will also yield similarly satisfactory estimates of unobserved periods of exposure.

4. The validity of the final equations selected as predictors will be established against three sources:

a. The 1/3 of the data, not used to derive the prediction constants, will be predicted. Besides serving as a check on the validity of the final method of estimation, the errors of estimate will be determined against that data.

b. The data collected during previous studies in this laboratory will be used in a similar fashion. The equations developed to predict plus the estimates of acceptable error derived from the remaining 1/3 of the data will be used to postdict results available since the beginning of the series of experiments.

c. It is recommended that subjects be exposed to two or three prolonged experimental sessions (up to 24 hours) and that the observed effects of these few sessions be evaluated against the predictions based on this method of analysis.