

THE INTRAFAMILIAL TRANSMISSION OF RHEUMATOID ARTHRITIS—I

DESIGN OF THE STUDY

SIDNEY COBB, M.D., M.P.H., ERNEST HARBURG, PH.D., JOYCE TABOR, M.A.,
PATRICIA HUNT, B.A., STANISLAV V. KASL, PH.D. and WILLIAM J. SCHULL, PH.D.

Institute for Social Research, Department of Psychology, Department of Human
Genetics, The University of Michigan, Ann Arbor, Michigan

(Received 21 December 1967; in revised form 26 August 1968; further revised 6 January 1969)

THE PURPOSE of this study is to examine the familial aspects of rheumatoid arthritis, both genetic and social. First, it aims to see if there is significant clustering in sibships and marital pairs, then it aims to see if the disease appears with undue frequency in persons with certain social and psychological characteristics and finally, it aims to see if persons in families with certain social characteristics are peculiarly subject to the disease. The first report describes the design of the study and discusses its suitability for the purposes listed above; subsequent papers will examine in detail the method for and the validation of the measurement of rheumatoid arthritis; the genetic findings, or rather the lack of evidence for a genetic effect; the perception that respondents have of their parents; the social and psychological correlates of the disease; the connecting links between status incongruence and rheumatoid arthritis; and finally the association of rheumatoid arthritis in the wife with peptic ulcer in the husband. A preliminary description of the study was presented at the conference on Epidemiology and Genetics of Chronic Diseases [1].

THE SAMPLE

The design of this study involves the taking of three extensive interviews with the members of 49 family clusters. An ideal cluster is composed of 7 members as follows:

1. The key person, or propositus
2. The key person's spouse
3. A sibling of the key person
4. The sibling's spouse, i.e. the key person's brother- or sister-in-law
5. A paternal cousin of the key person
6. A maternal cousin of the key person
7. A person unrelated to the key person.

*The authors are indebted to Dr. IVAN DUFF, Director of the University of Michigan Arthritis Clinic for help with the sample and with validation of the RA measure. The work was supported by the following grants from the U.S. Public Health Service:—AM-06928, K3-MH-16709 and K6-MH-21844. In addition, preliminary work on the development of measures in the area of anger and resentment was supported by the Rust Foundation and some generous citizens of Pittsburgh, Pennsylvania.

Reprint requests should be addressed to Dr. SIDNEY COBB, Survey Research Center, The University of Michigan, Ann Arbor, Michigan, 48106.

Of course, not all clusters were ideal. The variations in this regard are reviewed below and in the appendix to paper III. But first, in order to fully understand the structure of this study, it is important to understand how the key persons and the other members of the cluster were selected.

The key persons were selected from two sources. The first is a pair of national sample surveys, and the second is the University of Michigan Arthritis Clinic. The resulting subsamples are sometimes referred to as the National sample and the Clinic sample.

In order to select the *propositi* for the National sample, a set of questions about arthritis and certain demographic characteristics was included in surveys. These surveys, representing a random sample of the United States and drawn on the basis of a stratified area sampling technique [2], are conducted by the Survey Research Center for a variety of purposes at least four times a year. Such samples afford a unique opportunity to assess a given disease in the population, because they are free of many biases that exist when samples are drawn from more restricted geographical areas or from clinics. It was possible to add a few questions to two of these surveys at moderate cost, where national screening for the purpose of this study alone would have been prohibitively expensive. The questions included are known to be sensitive indices for the possible presence of rheumatoid arthritis, and the *propositi* were then selected from among the national sample respondents for having suggestive evidence of rheumatoid arthritis, for being white and married and for having an appropriate set of relatives from which to select the cluster members.

The key persons for the Clinic sample were chosen with the help of Dr. Ivan Duff, Head of the Arthritis Research Unit at The University of Michigan Hospital. Dr. Duff provided a list of clinic patients who were considered to be definite cases of rheumatoid arthritis. These patients were then given a screening interview to determine whether or not we could obtain from them an ideal family cluster.

TABLE 1. SOURCES OF LOSS FROM THE ORIGINAL SAMPLE TO THE FINAL SET OF CLUSTERS

	Total	National sample	Clinic sample
White married persons with arthritis	276	193	83
Number of active clusters	49	28	21
Number dropped	227	165	62
Dropped for incomplete cluster	129	81	48
Dropped for medical reasons	40	36	4
Refusal	22	21	1
Dropped for two of the above reasons	29	21	8
Dropped for other reasons	7	6	1

Table 1 shows that 193 persons were selected in the National sample screening and, of these, 28 or 15 per cent finally became key persons for clusters that were carried to the end of the study. In the clinic sample a somewhat larger proportion, 21/83 or 25 per cent, survived to be included. The reasons for dropping key persons are listed in the lower part of Table 1. Over half of those dropped were dropped because their families could not provide a complete cluster. The next most frequent reason for being dropped was medical. In most instances in the National sample

(33 out of 36), this was based on an estimate after the first interview that the person in question did not have a high probability of having rheumatoid arthritis. Surely there were some key persons excluded that really did have the disease just as there were some included who did not. In the Clinic sample only four were dropped for medical reasons. In two cases, the clinic diagnosis was revised after further investigation. Two others were rejected because of intercurrent disease of such severity that the series of interviews might have proven burdensome. When all the data were in and the final classifications were made, 43 of the 49 key persons were classified as having rheumatoid arthritis. (The 6 key persons not classified as having rheumatoid arthritis were retained in the study, together with other members of their cluster, because we had all the other data on them. In the data analysis, they were, of course, treated as respondents without rheumatoid arthritis.) An additional 23 persons classified as having the disease were not key persons but were other members of the family clusters.

Table 1 also gives the data on refusals. By refusal is meant refusal by the key person or multiple refusals among the cluster members. Almost all of the refusals were in the National sample, suggesting that the Clinic sample may have been somewhat preselected for cooperativeness. While on the subject of participation it is worth noting that of the 367 persons completing interview A, only 18 or 5 per cent refused to complete the succeeding 2 interviews. Only 1 cluster had to be dropped because of refusal after interview A. The additional losses due to the late medical rejections brought the final sample down to 324 persons. The distribution of these persons by sample and by position in the cluster is presented in Table 2.

TABLE 2. NUMBER OF INDIVIDUALS BY POSITION IN THE CLUSTER, SEX, AND SAMPLE

		Key person	KP's spouse	Sib	Sib's spouse	Paternal cousin	Maternal cousin	Unrelated	Total
National sample	M	6	22	14	13	8	13	21	97
	F	22	6	13	15	14	11	6	87
	T	28	28	27	28	22	24	27	184
Clinic sample	M	9	12	11	10	13	6	11	72
	F	12	9	10	11	6	12	8	68
	T	21	21	21	21	19	18	19	140
Total sample	M	15	34	25	23	21	19	32	169
	F	34	15	23	26	20	23	14	155
	T	49	49	48	49	41	42	46	324

Since it was originally contemplated that the principal analyses would compare those relatives who had rheumatoid arthritis with those who did not, the siblings and cousins were selected for being within 5 yr of age of the spouse of the key person so that there would be no more than 10 yr difference in age between the cluster members, and the differences in age between the propositus and the cluster members would be approximately constant. In an ideal cluster all members are married and living with spouse. A cluster was not accepted unless the propositus and sibling were living with spouse. We did retain 2 clusters with no married cousin and 1 cluster in which the sibling refused the last interview.

The unrelated person, or control, was also drawn from a national random sample and was required to be Caucasian, living with spouse, within 5 yr of age, and of the same sex as spouse of the propositus. In addition they were matched as closely as possible on family income, education, urban residence and region of the country.

DATA COLLECTION

Each of these persons was interviewed 3 times. Spouses of cousins and spouses of unrelateds were included on the third and final interview, as the content of this form was designed to obtain a description of the marital relationship. In so far as possible, the interviews were taken at 4-month intervals in order to eliminate the effects of the seasonal variation in the disease and to minimize the underidentification of persons with rheumatoid arthritis due to the remittant nature of the disease. This afforded an unusual opportunity for extensive data collection plus an opportunity in some instances to extend the inquiry on a given subject over more than 1 interview, thus providing a time perspective rather than just a momentary report of the present state.

As noted above, the technique for measuring rheumatoid arthritis and its validation will be presented in the next paper of this series. Here we will present the general nature of the variables examined and the general style of the inquiry. Each interview began with an enquiry about arthritis followed by DUNN's index of peptic ulcer [3]. Interview A contained demographic information about the individual and his parents and a variety of bits of health-related data. Interview B was concerned mostly with the perception of parents and certain mental health variables, particularly anger, aggression and depression. Interview C was devoted almost entirely to the nature of the marital interaction and related needs, personality characteristics and affective states.

The interviews were fully specified and all but a very few of the questions were of the closed end variety, that is either they required an answer codable in its own right, such as yes or no, or a date, or they had alternatives from which the respondents had to choose one. In order to present a varied and interesting interview the response scales varied in style from 4 to 7 points in length. Considerable pre-testing was necessary in order to establish a comfortable flow for the interview and to eliminate ambiguities. When each form was perfected a detailed instruction manual was prepared in order to minimize the variations between interviewers.

The interviews were all conducted by the well-trained and supervised national interview staff of the Survey Research Center [4]. They were usually taken at the home of the respondent though on occasion one was taken at a man's place of work. Of course, a great deal of travel was involved because in our mobile society many of the clusters were widely dispersed. In general a single interviewer took all 3 interviews with a given respondent and handled as many of the cluster members as were close enough geographically. Sometimes as in the case of the couple who wintered in Florida and summered in northern Michigan, 2 interviewers were involved for 1 pair of respondents. The interviewers reported that the respondents were substantially more cooperative and open than the respondents approached in the periodic economic surveys that they make.

ANALYSIS

When the interviews were returned, they were coded by the regular coding staff

of the Survey Research Center. This staff generally maintains its mean frequency of errors at or below 0.5 per cent for straightforward coding of this nature. From this operation resulted 27 decks of IBM cards and the data reduction began. Indices were constructed by combining items. The relevant details will be presented in the appropriate articles.

The structure of the final sample is presented in Table 2. Each of the 324 persons completed all 3 interviews. In addition, 70 of the spouses of the 83 cousins and 44 of the 46 spouses of the unrelated individuals were given interview C (they are not listed in Table 2). Lest the completion rate for cousins' spouses seem low, it should be noted that 10 cousins were accepted who were not married. This means that 96 per cent of the spouses accepted the interview.

At this point it is appropriate to examine a number of characteristics of the final sample and estimate the extent to which it might be considered representative of the U.S. population because this, of course, has bearing on the extent to which we can generalize from the findings. Table 3 presents a comparison of the National

TABLE 3. CHARACTERISTICS OF THE TWO SUBSAMPLES WITH COMPARISON TO NATIONAL NORMS WHERE APPROPRIATE

Variable	National norm for Caucasians born c. 1915	National sample	Clinic sample	Est. percentage of variance accounted for (%)
Mean age in years	..	49.6	49.8	—
Mean years of education	11.7 [6]	10.9	11.5	—
Mean family income	\$5800 [6]	\$4960	\$6630	8.8
Mean occupational level*	36.5 [7]	34.9	43.4	2.5
Perceived social class†	..	3.4	3.1	1.4
Size of sibship per individual	5.0 [8]	6.1	5.7	—
Percentage Roman Catholic	27.8 [6]	12.0	24.0	2.0‡
Percentage foreign born or with foreign born parents	11.4 [9, 10]	15.0	21.0	—
Percentage having lost a parent before age 14	20.0 [8]	12.0	11.0	—
Percentage previously married	17.0 [11]	17.0	18.0	—
Age at marriage; males	25.0 [12]	25.0	25.0	—
Age at marriage; females	22.0 [12]	22.0	22.0	—
No. of children per family	2.4 [13]	2.8	2.4	—
Mean length of time to complete interview	..	45 min	44 min	—

*Duncan Code: high number equals high prestige.

†One (1) equals upper class.

‡Estimated from critical ratio for difference in percentages.

and Clinic samples with each other and with national data where available. In the third column the per cent of variance accounted for is given if the *t*-test for the difference between the two samples is significant at the 0.05 level. As can be seen, the age for both groups is very close to 50. Of course, the men are somewhat older than the women as is usual in married pairs. With respect to education, income, occupation and perceived social class, the Clinic sample is significantly better off than the National sample. However, the relevant national averages for the age group lie between the two samples for income and occupation while for education the differences between the two samples and the national average are trivial. It should be noted that only for income is an appreciable proportion of variance accounted for by the

difference between the samples. (The estimated proportion of variance was calculated from the formula

$$\text{est } W^2 = \frac{t^2 - 1}{t^2 + N_1 + N_2 - 1}$$

as suggested by HAYS [5].) It seems that the National sample is below the national average on these social class variables which seems curious until we look at the next variable which is the size of the sibship. Clearly, in order to have living married siblings in the study, we have excluded sibships of a size one and increased the probability of selecting large sibships. Since larger families are found in the lower social classes, it is probable that there has been a selection in favor of the lower social classes. Happily, this is offset by the inclusion of the Clinic sample. Curiously, there are somewhat fewer Catholics and somewhat more foreign born or with foreign born parents than the national averages; but these variations are not very disturbing. The percentage having lost a parent before age 14 seems low. This again is probably related to the selection for large sibships. As a final comment, those who are confused by the size of the sibship per individual and the number of children per family will do well to consider 2 sibships of size 6 and 1 respectively. If they are averaged across individual, then:

$$\frac{6 \times 6 + 1 \times 1}{7} = 5.3;$$

if across families, then:

$$\frac{6 + 1}{2} = 3.5.$$

In addition to the examination by sample we have examined the same variables by sex and position in the cluster. The only striking finding is that male key persons from the National sample, who number only 6, stand out as rather strikingly different from all 28 other age, sex, and sample groups, including the male key persons from the Clinic sample, as being lowest in education, income and occupation and next to the lowest on perceived social class. Furthermore, they come from the largest sibships, are the oldest at time of marriage, and took the longest time to fill out the questionnaire. We conclude from this that our male key persons are not homogenous. However, only 3 of these 6 male key persons were eventually classified as rheumatoids. Conclusions about males with rheumatoid arthritis will be therefore more tentative. Furthermore, we are cautioned that male rheumatoids in attendance at clinics may be significantly different from those not attending clinics. This last argument is particularly cogent when it is realized that male rheumatoids in the U.S. National Health Examination Survey [14] are strikingly low on education and income, which suggests that it is probably the Clinic sample that is unusual rather than our National sample.

In passing it might be mentioned that there is an undue proportion of the siblings and the siblings' spouses who have been married more than once. The meaning of this finding is not immediately apparent.

DISCUSSION

The study was designed to test certain familial hypotheses. In addition, the material will be used for comparison of persons with and without rheumatoid

arthritis. It is therefore worthwhile to examine the extent to which the data are appropriate to these two purposes.

First, let us look at the problems of homogeneity. The material has been examined on a variety of demographic variables by sample, by sex, and by position in the cluster. The conclusion is that the material is really very homogeneous and that it is quite appropriate to combine samples. There are, however, two points of inhomogeneity—and since inhomogeneity may lead to spurious correlation [15], it is well to keep these points in mind. The first is the social class difference between the Clinic and National samples. This is not significant for education and sibship sizes and accounts for only a trivial proportion of the variance with respect to occupation and perceived social class. Only with respect to income does this dimension account for a proportion of the variance that might be large enough to give any concern at all. The second is the inhomogeneity of the male key persons and, therefore, to a large extent, of the male rheumatoids. This inhomogeneity is large though not statistically significant because of the small number of cases. Intuitively we feel that it is large enough so that conclusions with regard to rheumatoid arthritis in the male should be approached with caution. The conclusion that the material is suitably homogeneous with the cautions indicated is equally applicable to the family analysis and to the case-control analysis. Moreover, corrections for sample differences, where appropriate, and the very low magnitude of the correlations between the social class and the psychological variables, should further strengthen one's confidence in the results.

Next, it is important to estimate the extent to which it is appropriate to draw general conclusions from the sample in hand. It was noted above that the total sample is not appreciably different from the nation as a whole on certain social class and marriage variables. However, it is important to point out that the sample contains fewer Roman Catholics and more persons of foreign extraction than one would expect in a random sample of the nation. Because of the design of the study the sibship size is large and there are very few single people. Finally, there has been a selection in favor of the more cooperative, for the uncooperative refused to be interviewed.

In looking at those analyses in which rheumatoid arthritics are compared with others, the reader will do well to keep in mind the fact that the design involves a variety of linkages between the cluster members. Some of these involve the sharing of genes as in the blood relationships and others involve the sharing of the environment, e.g. the marital relationships. It is conceivable that these linkages or relationships may tend either to exaggerate or diminish estimates of differences that truly exist in the population.

In closing it is worth mentioning that persons with rheumatoid arthritis who are under continuous treatment may be appreciably different in their personality characteristics and their attitudes from rheumatoids who avoid or minimize their medical care. By virtue of the fact that 21 of the 49 key persons were drawn from the arthritis clinic the sample is somewhat overloaded with cases under continuous care. This matter will be discussed further in the next paper. Throughout this series of reports the reader can assume that differences between samples have been routinely looked for. If they are not mentioned, none were found.

SUMMARY

A sampling of 49 family clusters consisting of a key person with arthritis, his spouse, a sibling and the sibling's spouse, 2 cousins and an unrelated individual have been interviewed 3 times with regard to their arthritis and a variety of social and psychological factors. The sample has been drawn in part from a national random sample and in part from an arthritis clinic. The two subsamples have been found sufficiently homogeneous for combination and some of the strengths and limitations of the design have been discussed.

REFERENCES

1. COBB, S.: Intrafamilial Transmission of Rheumatoid Arthritis, *Genetics and the Epidemiology of Chronic Disease* (Eds. NEEL, G. V., SHAW, M. W. and SCHULL, W. J.). U.S. Government Printing Office, Washington, 1965.
2. KISH, L. and HESS, I.: The Survey Research Center's National Sample of Dwellings. Institute for Social Research, Ann Arbor, 1965.
3. DUNN, J. P.: *Duodenal Ulcer in Executives with Consideration of Methodological Problems in Formulating an Association Index*. Thesis, Graduate School of Public Health, University of Pittsburgh, 1959. (Available from University Microfilms, Ann Arbor, Michigan.)
4. SURVEY RESEARCH CENTER: *Interviewers Manual*. Institute for Social Research, Ann Arbor, 1966.
5. HAYS, W. L.: *Statistics for Psychologists*. Holt, Rinehart and Winston, Chicago, 1964.
6. U.S. BUREAU OF CENSUS: *Statistical Abstract of the United States: 1960*. (81st edition.) Washington, 1960.
7. KATONA, G., MUELLER, E., SCHMIEDESKAMP, J. and SONQUIST, J.: 1966 Survey of Consumer Finances. Survey Research Center, Mon. No. 44. Institute for Social Research, Ann Arbor, 1967.
8. CHEN, E. and COBB, S.: Family structure in relation to health and disease. *J. Chron. Dis.* **12**, 544, 1960.
9. U.S. BUREAU OF CENSUS: *U.S. Census of Population: 1960. Subject Reports. Detailed Characteristics—U.S. Summary*. Final Report PC (2)—4E. U.S. Government Printing Office, Washington, 1966.
10. TAEUBER, T. and TAEUBER, I.: *The Changing Population of the United States*. Wiley, New York, 1958.
11. U.S. BUREAU OF CENSUS: *U.S. Census of Population: 1960. Subject Reports. Marital Status*. Final Report PC (2)—4E. U.S. Government Printing Office, Washington, 1966.
12. U.S. BUREAU OF CENSUS: *U.S. Census of Population: 1960. Subject Reports. Age at First Marriage*. Final Report PC (2)—4E. U.S. Government Printing Office, Washington, 1966.
13. U.S. BUREAU OF CENSUS: *U.S. Census of Population: 1960. Subject Reports. Women by Number of Children Ever Born*. Final Report PC (2)—4E. U.S. Government Printing Office, Washington, 1966.
14. ENGLE, A.: personal communication.
15. BATSCHLET, E.: Spurious correlation of the age of onset, with special reference to asthma bronchiale and atopic rhinitis. *Biomet. Z.* **4**, 111, 1962.