

AGENCY FOR INTERNATIONAL DEVELOPMENT  
 WASHINGTON, D. C. 20523  
**BIBLIOGRAPHIC INPUT SHEET**

FOR AID USE ONLY  
*Batch 92* ARDA

1. SUBJECT CLASSIFICATION	A. PRIMARY Science and technology	T000-0000-0000
	B. SECONDARY Applications	

2. TITLE AND SUBTITLE  
 Socio-economic analysis of impact of water projects on schistosomiasis; preliminary report

3. AUTHOR(S)  
 Rosenfield, P.L.; Gestrin, P.J.

4. DOCUMENT DATE 1978	5. NUMBER OF PAGES <del>100</del> 105 p.	6. ARC NUMBER ARC
--------------------------	---	----------------------

7. REFERENCE ORGANIZATION NAME AND ADDRESS  
 Resources

8. SUPPLEMENTARY NOTES (*Sponsoring Organization, Publishers, Availability*)

9. ABSTRACT  
 Describes progress made on an AID-supported socio-economic study to analyze and control the incidence of schistosomiasis, a parasitic disease transmitted by snails. The project seeks to help decision makers safeguard against the increased danger of schistosomiasis caused by large-scale water projects. In this report, a previously developed transmission model is refined, using data from the WHO/Tanzania schistosomiasis pilot control project. Modifications include refining the incidence rate equation to include egg counts and human water contact, and designing a mechanism to account for population movement. The study provides numerous figures and tables, information on data processing methods calculation procedures for the terms in the model, criteria for choosing terms, and a suggested statistical system for schistosomiasis control projects.

10. CONTROL NUMBER PN-AAG-056	11. PRICE OF DOCUMENT
12. DESCRIPTORS Economic analysis Environmental factors Models Schistosomiasis	13. PROJECT NUMBER 931113300
	14. CONTRACT NUMBER AID/ta-C-1465 Res.
	15. TYPE OF DOCUMENT
Tanzania Water resources Water supply	

Return

AID/TS-C-1465 Rec  
Resources  
ARDA PI-AAS-55

SOCIO-ECONOMIC ANALYSIS OF IMPACT OF  
WATER PROJECTS ON SCHISTOSOMIASIS

Preliminary Report

AID-Contract No. 931-1133

**RESOURCES FOR THE FUTURE, INC**

1755 MASSACHUSETTS AVENUE, N.W.

WASHINGTON, D. C. 20036

SOCIO-ECONOMIC ANALYSIS OF IMPACT OF  
WATER PROJECTS ON SCHISTOSOMIASIS

Preliminary Report

AID-Contract No. 931-1133

by

Patricia L. Rosenfield, Ph.D.

and

Phyllis J. Gestrin, Ph.D.

Resources for the Future  
Washington, D. C.

July 1978

## TABLE OF CONTENTS

Text of Report.....	1
References.....	30
Figures and Tables.....	33
Appendix I-Data Processing.....	I-1
Appendix II-Calculations.....	II-1
Appendix III-Statistical System.....	III-1
Appendix IV-Consultations.....	IV-1

SOCIO-ECONOMIC ANALYSIS OF IMPACT OF WATER PROJECT  
ON SCHISTOSOMIASIS: PRELIMINARY REPORT TO AID FOR  
PROJECT NO. 931-1133

by

Patricia L. Rosenfield  
Principal Investigator

and

Phyllis Gestrin

Introduction

The purpose of this report is to describe progress made towards completion of the objectives of the AID-supported study, Socio-economic Analysis of Impact of Water Projects on Schistosomiasis. A brief report of consultations with researchers at WHO and the Ross Institute and a preliminary version of a protocol for data presentation in future schistosomiasis control projects are appended.

The objective of this stage of the project is to refine a previously developed schistosomiasis transmission model (1) by use of data from the WHO/Tanzania Schistosomiasis Pilot Control Project, Mwanza District (Misungwi), Tanzania (hereafter called Tanzania 001). The following modifications to the model to meet Objective I (2) were considered: revising the incidence rate equation to include egg counts; revising the incidence rate equation to include human water contact or an alternative measure of accessible snail habitats; incorporating a random component in the model to account for population movement; incorporating terms in model to

account for the distribution of worms in the human population; and incorporating age and sex differences in the model. The relative importance of these different terms in estimating incidence rates and then predicting prevalence or egg count levels in the project population are also examined.<sup>1</sup> Later reports will describe: a) the results of cost-effectiveness analyses of control measures (Objective II); and b) the results of integrating Objectives I and II into a practical, tested methodology for predicting the impact of water resources development on schistosomiasis transmission (Objective III).

This report is organized as follows: 1) discussion of the transmission model and previous work with it; 2) description of the Tanzania 001 project; 3) discussion of the data processing techniques used for our computer analyses; 4) preliminary results of modifying the transmission model with different variables by use of the Tanzania 001 project data; and 5) four appendices which mainly cover Tanzania 001 data presentation in general.

#### Schistosomiasis Transmission Model Development

The aim of the modeling efforts described here is to provide decision-makers with a methodology for predicting the impact of environmental changes on the level of schistosomiasis in the project population. With

---

<sup>1</sup>Incidence is the rate at which uninfected persons become infected over a given period of time. Prevalence is the number or fraction of infected persons in a population at any point in time. Egg counts are thought to be a measure of severity of the infection and are often called intensity of infection. Reversion is the rate of infected persons spontaneously (not due to treatment) becoming negative over a given period of time. We use the word "infection" to describe the fact that there are positive cases of schistosomiasis in a population. The word "disease" is reserved for use when clinical symptoms from the infection are observed

a methodology based on fundamental schistosomiasis transmission variables, control or preventive measures may be examined for their effectiveness in minimizing the adverse human health impacts of the anticipated environmental changes.

Environmental changes pertinent to schistosomiasis transmission result from the channeling of water in irrigation projects for food production or the impounding of water in reservoirs behind dams for power production or water storage. These changes may be large or small scale. The result of such changes is to make the area more desirable for human settlement thus attracting migrants who frequently bring schistosomiasis and its snail host with them. In addition, the increased amount of water in an area provides increased habitats for the snail intermediate hosts of schistosomes. Without adequate social services such as water supplies and sanitation facilities for the human settlements, inhabitants have no choice but to use snail-infested waters for a variety of domestic and economic uses. Thus, environmental changes frequently required for development purposes may result in lowering the quality of life for persons in the project area.

In many instances, the environmental changes occur where schistosomiasis is a long-standing endemic problem. The project then results in such rapid egg count increases or prevalence levels so as to create a serious disease problem. Many water development projects, however, are being built in areas where schistosomiasis has been unknown (for example in arid regions); resulting schistosomiasis transmission may then reach epidemic conditions. In either case, the project population suffers as

a result of good intentions. It is hoped that the predictive methodology described here can be used by project planners in alleviating or, more importantly, in preventing the adverse health consequences of large-scale water projects.

Before describing the methodology, it should be mentioned that adverse impacts are no longer made unwittingly. The relationship between schistosomiasis transmission and water resource<sup>s</sup> projects has been referred to often (3). With increasing demands for food and power in the developing world, it is now imperative that this long-recognized relationship be considered in water resource project design, construction, and operation.<sup>2</sup>

#### Modeling Objectives and Assumptions

The type and degree of sophistication of the methodology used are directly related to specific objectives. In this case, the main objective is to develop a predictive methodology to assist with development project planning and control program management. The objective is not to quantify all of the relationships in the schistosome life cycle but rather to use the minimum number of fundamental elements needed to predict changes in schistosomiasis transmission over time. These fundamental elements, infected persons and snail habitats, are directly affected by development projects and control measures. By relating the elements in a predictive model, one can estimate changes in incidence rates and prevalence in the project populations produced by the project. The

---

<sup>2</sup>This topic will be expanded in the final report with a discussion of donor agency activities to reduce the adverse environmental health impacts of water resources projects.

focus of the modeling effort described here is on the human population, since "evaluation of the efficiency of the control programs must be made in terms of its impact on human schistosome infection" (4). Snails, of course, play a critical role in transmitting infection, but it is the reduction of human suffering which concerns health and development planners.

The desire to keep the variables in the model to a minimum may become more understandable when one considers in detail the variety of measures of human infection and snail populations. Human infection encompasses variables such as egg counts, prevalence, incidence, reversion, and immunity. Snail population information covers infected snails, population density measures, habitat characteristics, and climatic variation.

Many medical researchers believe that human infection should be expressed in units of eggs passed. Egg output, however, "should be determined by a constant, reliable, and quantitative technique which must not vary from year to year" (5). This is often not the case and even more importantly, the reasons for variance in eggs passed from a given individual from year to year are not yet understood. The variance leads to complications in the statistical handling of egg counts, requiring transformation of the values to either the arithmetic mean or geometric mean for each age group. Yet, egg counts are a measure of intensity of infection and thus can assist in estimates of severity of infection and changes in transmission intensity. For these reasons, egg counts must be considered in the modeling attempt, if only to show whether or not estimates of control measure effectiveness vary when one uses egg counts or prevalence levels.

The second measure of human infection is prevalence. Prevalence levels, the proportion of the population or age group infected, are thought to be insensitive to short-term changes as a result of control programs but do reflect changes over the long-term. Prevalence is, however, the easiest variable to measure. Moreover, its denominator is the population size (or sample of same); therefore, it is a term comprehensible to planners and compatible with project analyses. Prevalence provides a useful, simple term for demonstrating project-induced changes.

Incidence, the rate at which uninfected persons become infected, provides a time dimension to transmission. Although difficult to observe without a carefully controlled study, incidence rates, like egg counts, provide a sensitive measure of changes in transmission. In addition, in the model we are interested in predicting changes of infection levels over time. Incidence rates reflect changes in transmission over time.

Two aspects of human loss of infection (besides losses exogenously induced by chemotherapy) need also to be considered. There is a spontaneous loss of infection resulting from natural death of the worm in the human body. S. haematobium is thought to have an exceedingly short life span, under three years, and light human infection with no re-exposure could die out without any treatment. S. mansoni and S. japonicum live longer in man; in extreme cases, S. mansoni worms have been found in infected immigrants in New York City and California 20 years after their arrival (6). Any modeling attempt should account for the worm death rate.

The second loss over time is due to immunity. The process of immunity to new schistosome infections is still not understood. It was thought to occur because observations of the age prevalence curves for schistosomiasis infections, especially S. haematobium, repeatedly show a decline after ages 15 through 25. The reasons for this decline have been the subject of debate (7). It is also generally agreed that natural immunity to schistosomiasis does not exist; however, immunity acquired from continual re-infection or high levels of infection is thought to be a significant possibility. Acquired immunity has not yet been identified in the field, for one must separate it from changes in water contact patterns, seasonal changes in egg output patterns, and other variables which might account for changes in the shape of the age-prevalence or age-egg output curves (8).

The snail as the intermediate host in schistosomiasis transmission has often been thought of as the weak link in the life-cycle and highly susceptible to control measures (9). The validity of this viewpoint is in doubt given the lack of success in controlling snail populations in any but the most limited environment. So many factors influence snail populations that a separate model is needed to account for the snail's role in transmission. Yet collecting information on food, light, density, and other other requirements is time consuming, and may not be feasible as part of the planning of large scale water projects. It may be sufficient to know which bodies of water are likely to harbor snails, which schistosome-bearing species of snails are in the area, and what role seasonal variation may play. In most areas, only a small proportion of snails are ever infected (10). One may be able to assume that any habitat

with the appropriate snail species which is used by humans is a likely transmission site. The knowledge of habitat dimensions and human water contact patterns may provide the requisite amount of information needed for predictive modeling. This assumption was the basis of earlier work and is examined here in the section on model modifications.

The assumptions about which data to use are therefore the following:

1) egg counts would be most desirable but if not available, prevalence levels are sufficient as surrogate measures of human infection; 2) incidence rate is critical since it is assumed to be a sensitive measure of transmission changes; 3) immunity, not yet a proven factor in transmission, could be indirectly reflected through egg count or prevalence changes; 4) reversion rates have been calculated from the egg count data and are assumed to be age-specific; and 5) snail habitats that are transmission sites or estimates of human water contact frequency are satisfactory surrogates for infected snail population data.

Implicit throughout the above discussion is the desire to maintain simplicity in the model and its data requirements. Many diseases can be appropriately modeled by the simple catalytic expression developed by Muench (11).

The expression that we chose to use straightforwardly defines the rate in change in prevalence levels (y) over time (t) as a function of uninfected persons become infected at some rate (A) and infected persons losing the infection at some rate (B):

$$\frac{dy}{dt} = A(1 - y) - B(y) \quad [1]$$

This equation may be solved to give the following difference equation (12):

$$y_{t + \Delta t} = (y_t - \frac{A}{A + B})e^{- (A + B) \Delta t} + \frac{A}{A + B} \quad [2]$$

This equation is used to predict changes in levels of infection from one time period (t) to the next (t + Δt).

As mentioned above, it is assumed that the incidence rate (A) is a function of infected persons or egg output, and snail habitats or human water contact. Incidence rates may be estimated by use of regression analyses if field observations are available.<sup>3</sup> Regression analysis results indicate the relative importance of the independent variables in estimating the variance in incidence. Although the relative importance of the independent variables in predicting incidence rates may change when applied to a new site, the same independent variables may still be appropriate. In addition, the form of the equation relating them may remain constant. To test the hypothesis of generality of the variables and equation form, the results from this study were compared with two previous analyses. These earlier studies indicated that the most significant form of the equation was the non-linear product of an environmental/behavioral term (H) times an epidemiological term (P):

$$A = \beta_0 (H_t^{\beta_1} \times P_t^{\beta_1}) \quad [3]$$

where A was the incidence rate, t the time units, and β<sub>0</sub>, β<sub>1</sub>, β<sub>2</sub> estimated

<sup>3</sup>Incidence rate studies are carried out as follows: A group of a particular age or range of ages is determined to be definitely negative by repeated urine or feces examination. This group, with no replacement for drop-outs, is followed over time to determine who becomes positive after, for example, one year. The same group minus the positive is followed and examined the next year. The study may continue as long as the sample size is sufficiently large. The incidence rate is measured by dividing the new positives in the next year by the total number of negatives the year before.

by the regression analyses. In the original study in Iran, H was measured as meters of accessible snail habitat per village per year and P was the number of infected persons per village per year. In the next study in St. Lucia, H was the human water contact parameter per age group (a measure of frequency and duration of human water contact with snail habitat) and P was the number of infected persons per age group.

The expression for the loss rate B has been assumed to be related to the natural death rate of the worm, physiological characteristics of the human host and immunity. Two studies have attempted to observe loss rates in the field (B).<sup>4</sup> These results compare favorably with mathematical estimates (obtained by solving equation [2] for B). A general estimate of  $B = 0.2$  was used in Iran and age-specific estimates of  $B = .3$  for 0 to 9-year-olds and  $B = .2$  for older ages were used in St. Lucia. Sensitivity analyses show this term to be highly important in prevalence predictions. Site-specific field studies would assist in improving its usefulness and reliability for modeling.

The results of using the model in Iran and St. Lucia are given in Figure 1. The estimates in Iran were made on a village basis; the estimates in St. Lucia on an age-specific basis.

---

<sup>4</sup>The loss rate or reversion rate is measured in the field in a manner similar to incidence studies. A group of positives is followed over time, usually one year, to determine the number who lose the infection. The positive group must not be treated so that the loss may be attributed to natural conditions.

### Schistosomiasis in Tanzania

Schistosomiasis has been studied in Tanzania since the 1950's with the East African Medical Research Institute, Mwanza acting as the focal point for the research (see Figure 2). Researchers at the institute have measured prevalence levels of the two species in Tanzania, S. haematobium and S. mansoni. They have also instituted control activities, the most extensive one being the project under discussion (14). In addition, there has been a number of programs studying transmission specifically in large-scale irrigated areas, with the Arusha Chini scheme in the North as the focus for control, economic, and epidemiological studies (15).

S. haematobium is found throughout Tanzania, with prevalence levels as high as 80 percent common to the north and northeast of Lake Victoria (16). Although S. mansoni is less generally widespread in Tanzania, prevalence levels may be as high as 50 percent around Lake Victoria (17). A variety of snail hosts are found in the country. Ideal habitats occur in Lake Victoria, irrigation systems, small ponds, dams, and permanent and seasonal water courses (18).

The public health importance of schistosomiasis infections in Tanzania has been summarized by Jordan in 1966 (19). He described several studies which showed a variety of severe clinical effects from the infection, especially in children. Widespread urological changes for S. haematobium were identified. S. mansoni infection at that time appeared to be less severe in Tanzania than in other parts of the world.

In the 1970's, Tanzanian development has been influenced by President Julius Nyerere's villagization program, especially by the establishment of

Ujamaa villages which stresses cooperative activity. (20). The movement and settlement of large numbers of persons is bound to have a dramatic impact on schistosomiasis transmission, especially since the expected improvements in sanitary conditions (by the provision of water supplies, latrines, and health education) have not yet taken place (21). Moreover, small-scale irrigation projects continue to play a role in providing water for crop production. Schistosomiasis prevalence may be spreading as a result of too-limited control activities and environmental changes brought about by new living patterns.

#### Schistosomiasis in Misungwi, Tanzania

The data for this study comes from the World Health Organization/Tanzania Schistosomiasis Pilot Control and Training Project, Mwanza District, Tanzania, run by the East African Medical Research Institute under the direction of D. V. M. Eyakuze, National Project Director, and Dr. Fergus McCullough, WHO Project Leader. The project began in April 1967 and ended in December 1973. Two contrasting areas were chosen for the project site, an urban situation at Mwanza and a rural area around Misungwi (Figure 3). The aim of the project was "to determine the feasibility of schistosomiasis control within present resources, and to evolve methods which may be applicable, with suitable modifications, to other parts of East Africa" (22). In Misungwi, the rural area, S. haematobium infections predominated, "with an average prevalence of over 60 percent" (23). The major snail host was B.(P) nasutus which was found in "small man-made water bodies" (24). In Mwanza, the urban area, S. mansoni was the main species of the parasite, with prevalence

levels reaching 30 percent (25). The intermediate hosts, various Bio-  
mphalaria snail species, were found "in streams, lakeshore waterbodies  
and Lake Victoria itself" (26).

The Misungwi situation, like much of this part of Tanzania, repre-  
sents a "complex transmission pattern" (27) where in a relatively uni-  
form environment, "the infection, in both the definitive and intermedi-  
ate hosts, is characteristically widespread and relatively non-focal"  
(28). The pilot area was to the northwest of Misungwi Town covering an  
area of 76 km<sup>2</sup> and a population of 4000 (see Figure 4). Although  
S. haematobium was the major species present, pockets of S. mansoni  
infections did occur with prevalence levels of 10 percent.

The Misungwi area was chosen as the focus for this study because  
the data were complete, long term, and also were recorded on punch  
cards so that a computer tape could be easily made. The area was  
divided into five sectors so that different control measures could be  
tested in different sectors, thus providing a basis for comparison of  
control measure effectiveness. In May 1976, Dr. Rosenfield with assist-  
ance from the pertinent staff at WHO was able to make such a tape. In  
addition, she had generous access to all the project reports and other  
unpublished material. In May 1978, a second visit to WHO enabled  
Drs. Rosenfield and Gestrin to complete the necessary data collection  
and to confer with Dr. McCullough about the progress of the study to  
date. Consultations with Dr. McCullough, Dr. Bradley in London, and  
others are described in Appendix IV.

The work plan and results of the Project in Misungwi were stated  
as follows in the Tanzania 001 Final Report:

At Misungwi the work was planned in three interconnecting phases: pre-control, control and assessment. During the baseline phase detailed information was obtained on population, prevalence, snail hosts and their habitats. The pilot area was divided into five sectors. In Sectors I to IV, those habitats which were found to be important transmission sites were molluscicided with three applications of Bayluscide (1 p.p.m.) per year. In Sectors II and III mass chemotherapy using niridazole, in addition to the routine mollusciciding operations, was given. Sector V was used for comparative purposes. The pre-control phase, which began in July 1967, was completed in April 1970; mollusciciding was started in May 1970 and completed in June 1972; and mass chemotherapy was given in July 1970. The results were very encouraging, and indicated that S. haematobium infection was substantially reduced by the control measures undertaken at Misungwi (29).

The project data sources available to us included the following:

1) project quarterly and annual reports from 1967 to 1973; 2) published reports on the project; 3) computerized data on the human population and snail habitats; and 4) maps indicating habitat and household locations. The data themselves covered a wide range of variables.

For each individual the following data were obtained: study number; sector number; household number; number living in household; type of habitat used for bathing/washing or swimming; number examined in household; individual number; sex, age; length of stay in area; religion; previous treatment for schistosomiasis; education; 1968, 1970, 1972 urine examination results, such as pre- and post-treatment egg output (eggs per 10 ml.); head of household; year of registration; population movements (in and out of area or sector, or movement within sector); year of movements; and treatment history.

Snail habitat information included the following: study number; habitat number; date of survey; type of habitat; character of habitat; vegetation and vegetation clearance; measurements when full; frequency of water use; type of water use; snails found in habitat; number of houses within 300 yards of habitat; transmission potential; area and survey number.

Map data included information on the project area as a whole and on the separate sectors.\* For the project area, there were maps showing the location of households and different types of habitats (dams, ponds, banded fields) without reference numbers, and maps of snail species distribution. One map of Sector IV had households and habitats identified by number. Another sector map gave household numbers for Sector I with habitats shown but not numbered. A third sector map gave habitat numbers for Sectors II, III and IV and showed households without household numbers. In addition, there were more detailed maps for the Masawe indicator area, a section set aside in Sector II for special studies and as a statistical control area. The Mitando Chemotherapy area in Sector II was also mapped carefully. In both these special areas numbers were not given for household or habitats.

As may be seen from the above listing, the data available were extensive. The project determined that the age-sex structure for the sectors were similar and overall compared favorably to that for rural Tanzania (30). (See Figure 5.). Indeed, within the five sectors, transmission conditions were fairly uniform (31). Household distributions were similar, as were water contact patterns and socio-economic conditions. Studies indicated that

age-specific

\* After writing this report, we learned that Dr. David Bradley had received from Mwanza a more complete series of project maps. We have not yet had the opportunity to examine those maps.

prevalence and age-specific egg output were similar. The distribution of snail habitat types varied, with some sectors having more dams. Schistosome-bearing snail species, however, were fairly evenly distributed among the sectors.

Although population mobility was high, much of the movement was within a radius of 100 km. Thus, "S. haematobium infection at Misungwi is very highly endemic and perhaps its most remarkable feature is its homogenous, relatively non-focal distribution which stems from the widespread distribution of the snail hosts. . . and evenly scattered populace. . ." (32).

Control Operations (Will be discussed in next report.)

#### Data Processing

In reviewing the data for modeling purposes, a number of concerns arose. It should be emphasized that modeling schistosomiasis transmission was not one of the objectives of Tanzania 001. Because of that, we had to reorganize the data as given. In addition, we had questions regarding computer coding of data. In Appendix I, we describe the steps necessary for the data analysis and the problems encountered. In Appendix II, we give the calculations we used for putting the data into usable forms. As a result of our work with this project and others, we have included in Appendix III a possible form for reporting data which would facilitate project evaluation and transmission modeling studies.

The most critical gap in the data is the difficulty in comparing houses and habitats. Water use patterns are now accepted as crucial indicators of transmission (33). It would be highly desirable to know

which household used which habitats or which habitats served which household. With the background material and data available to us, it was possible to make some assumptions about these patterns for Sector IV (the only sectors where households and habitats were both mapped and numbered). In the other sectors, more assumptions will be necessary to make such connections. For this reason, Sector IV served as the basis for the present study.(see Figure 6).

In Sector IV, we considered different assumptions about the distance walked by individuals to use the different water sites. The sector, approximately 13.4 square miles, contained 121 households, 660 persons, and 102 habitats divided into three sections: Ibilibishi, Igokelo, and Nange. Information was not available on which persons used which habitat for a given purpose, duration, or quantitative frequency. Thus, we decided to use as the habitat variable one comparable to that in the Iran study, feet of accessible snail habitats.

We considered two different assumptions of accessibility. The first was that people did not move too far from the household for the bulk of domestic activity and that they had access to all habitats within a radius of one-eighth mile. The second was that movement was really more extensive and that an outer limit of one-half mile radius would be more accurate. In rural areas, people do walk great distances to obtain water. In fact, in some parts of Sukumaland people walk nine miles a day to water their cattle (34). However, in this particular sector, almost every household was close to a number of different habitats, and it is possible that most of the water contact

took place close to home. In fact, the project staff did record number of households within 300 yards of each habitat, an indication that this assumption might be acceptable.

Another problem with using the data from the maps was that topographic information was not easily usable (only edges of valleys were mapped without altitude contour lines). We assumed the area was one of gently rolling hills and, except where noted on the maps, the hills offered no obstacles to habitat accessibility.

### Regression Analysis

In order to run the regression analysis for the model, it is necessary to have incidence rate values. In this study, incidence rates were measured for each age from two through nine years over a 2-year interval. For the initial regression analyses, we therefore used 2-year incidence rate values for the two through nine-year-olds. To use the overall model for these ages, it was necessary to estimate a separate population growth rate for this age group. Information given by Ruysenaars, et al. (35) indicated that the annual population growth rate in Sector II for the two to nine-year-olds was three-fourths of a percent. Working with the data from Sector IV, we also estimated an annual population growth rate for the two to nine-year-olds of three-fourths of a percent. We used this as an estimate for annual population increases. In addition, we experimented with different groupings of age-specific data to see if groupings changed results. The groupings we tried were 1) each age separately, 2) two through four and five through nine-year-olds similar to standard demographic tables, and

3) two, three through four, five through eight, and nine-year-olds, a grouping warranted by the prevalence and incidence data.

As has been discussed above, the purpose of the regression analyses is to estimate incidence rate values for use in the transmission equation [2]. Incidence rates have been assumed to be a function of the non-linear interaction of environmental (snail habitat), behavioral (water contact and epidemiological (number positive, egg counts) terms. With the Tanzania data, it has been possible to experiment with a number of different formulations based on environmental and epidemiological assumptions. We could only indirectly include the behavioral term because no separate quantitative study had been carried out on human water contact patterns. We separated the data by age as described in Appendix I. The following age-specific combinations were examined:

Incidence rate as a function of:

1. Volume (in cubic feet) of accessible snail habitats at the given distance (miles) from each household per age group ( $V_i$ ) times number of positive persons per age group ( $P_i$ )
2. Perimeter (in feet) of accessible snail habitats at the given distance (miles) from each household per age group ( $PM_i$ ) times number of positive persons per age group ( $P_i$ )
3.  $V_i$  times arithmetic mean egg counts per age group ( $E_i$ )
4.  $PM_i$  times arithmetic mean egg counts per age group ( $E_i$ )
5.  $V_i$  times geometric mean egg counts per age group ( $GE_i$ )
6.  $PM_i$  times geometric mean egg counts per age group ( $GE_i$ )

where  $i$  indicates age. Regressions one and two were run for one-half mile habitat estimates. All were run for the one-eighth mile estimates. The methods for calculating the different terms are given in Appendix II along with the decision criteria for accessible snail habitats.

The regression results are given in Table 1, and the comparison of predicted versus observed incidence graphed in Figure 7 for one-eighth and one-half miles. As mentioned before, these are the results for ages two through nine for 2-year incidence. The results indicate that the equations using perimeter and number positive, and perimeter and arithmetic mean egg counts are the most significant estimates of incidence. However, equations containing volume terms predict incidence almost as well as those with perimeter terms. This is probably to be expected since volume is a function of perimeter. More surprising is the absence of a difference between one-eighth and one-half mile results. This indicates that people probably make frequent use of habitats over as great a distance as one-half mile.

The choice of which equation to use to predict incidence in the model (see equation [3]) can be determined by how one wants to use the model. Since for the moment we are not able to use egg counts in the recursive model, we chose to use the number positive equation for testing the model. To see if any differences (however unlikely) would result from using volume and perimeter terms, we tried equations 1 and 2 in separate testings of the model (see page 19).

To run the model, we need the following age-specific input information: baseline population (1968); baseline prevalence; volume or perimeter of habitat; and the reversion rate. Since in this preliminary

version of the model the effects of control measures were ignored, the values for the habitat term remained constant over the years of analysis. The size of each age group was assumed to increase by three-fourths of a percent every year. Later analyses will include the effect of in-and out-migration on the population size. The model was run at two-year intervals since the incidence data were 2-year incidence values. The output for the model may be thought of as the prevalence levels for each age group at the end of the 2-year time interval. In the future, the data analysis will be modified to reflect the movement of individuals to successively older age groups.

The results of running the model are graphed in Figure 8 where the predicted prevalence results are compared with the observed values. Running the model with the one-eighth mile and one-half mile regression equations and with volume and perimeter equations showed only slight changes. We believe this means that further work is necessary to obtain a better definition of the snail habitat term. It may be, however, that the snail habitat term is fairly robust; that is, changes in its value (as long as appropriately related to household) will not cause serious changes in prevalence predictions. In reality, if this is the case, then more effort will be given to modifying the epidemiological term.

#### Modifying the Epidemiology Term

The availability of the Tanzania 001 data enables the epidemiological term in the incidence rate prediction to be modified in a number of important ways. We are in the process of exploring modifications and shall report the results in a later paper. In this report, the modifications are discussed. The model is now a deterministic one: that

is, the effects of new infections introduced by migrants to the area cannot be predicted by the model. The Tanzania 001 data include information on population movements; therefore, a random component can be added to the model and tested with field data. In addition, the egg count data are untransformed numbers of eggs passed per 10 ml. of urine for each individual. This means that we are able to test a variety of statistical transformations of egg counts for modeling purposes.

There has been some discussion in the literature (36) as to the appropriate expression for the distribution of egg counts. Egg counts are an indication of the paired worm load in an individual. A person cannot pass eggs unless he has at least one male and one female worm that are paired. Moreover, a person will not be counted in a prevalence study unless he passes eggs, the usual indication of infection. Serologic tests to demonstrate infection are not widely used nor considered reliable. It would be helpful to have relationships linking prevalence, egg counts, and worm load. These relationships may be described by making statistical assumptions as to their distribution in a population. The critical variable, however, for describing the distributions is the worm load in the human body, a variable for which field data on humans are impossible to collect. Data from animal studies and human autopsies have been used as the starting point for estimating worm load distributions in a living human population. It is difficult to use these estimates in any serious quantitative way.

Macdonald was the first to assume a statistical distribution of worm pairs (37). He considered worm pairs to be distributed randomly in a human population and chose the Poisson distribution to describe real situations. The Poisson distribution is expressed as follows:

$$P(i/m, \infty) = e^{-m} m^i / i! \quad [4]$$

where  $i$  is the number of objects,  $m$  is the mean worm load, and  $\infty$  is the limit which  $k$ , the clumping parameter, approaches. This distribution has two requirements: 1) each case is independent from the next; and 2) the probability of success in finding a positive case is constant. The population of worms, eggs, and infected persons need to be independently, randomly distributed to fit the requirements for use of the Poisson distribution.

With animal and autopsy data measured after Macdonald's pioneering work, it was observed that worms were distributed in a clumped, not independently random, fashion. Together with the assumption that human water contact does not tend toward randomness, it has been assumed that the distribution of schistosomes is not random but aggregated or clumped (38). The clumping affects the fraction of eggs expected to be passed and therefore influences how one would model transmission processes. The distribution chosen to describe this clumped status is the negative binomial distribution:

$$P(i/m, k) = (1 - \alpha)^k \left[ \frac{(k-1)! (i+k)}{(k-1)! (k)i!} \right] \alpha^i \quad [5]$$

where  $i$  is the number of objects,  $m$  is the mean worm load,  $\alpha$  is  $m/(m+k)$  and  $k$  is the clumping parameter (39). The smaller the value of  $k$ , the

more clumped the worms are. As  $k$  decreases, the variance increases, and the worm load in a population may be characterized by most people having zero worms and a few people having very many worms. In the case of a Poisson distribution, the clumping parameter,  $k$ , approaches infinity (40).

The negative binomial may be appropriate to use when considering an individual's egg production over time. The Poisson may be more useful when looking at a population as a whole and trying to estimate the mean and variance as a function of the population's characteristics (Kerry Smith, RFF, personal communication, 1978). We have attempted with the Tanzania data to decide which distribution is more appropriate to use in that situation and if it is possible to use other distributions to relate egg counts to prevalence.

The use of egg counts requires changing the model to preserve its recursive nature. In the present version of the model, the number of infected persons is a term in the equation predicting incidence rate [3], and incidence rate is used to predict prevalence for the next year [2]. The next cycle of the model is started by computing number of infected persons from this prevalence prediction. To maintain the recursive nature of the model and make use of egg counts, it may be possible to relate egg counts and prevalence through human worm load. We described earlier the theoretical aspects of using egg counts. Here we relate that discussion to our study in Tanzania.

Bradley and May (41) have derived equations for these epidemiological relationships that we are attempting to use in the model:

if  $k \rightarrow \infty$  (Poisson distribution, male and female worms distributed together.)

$$y_t(m, \infty) = (1 - e^{-m/2})^2 \quad [6]$$

if  $k \rightarrow 1$  (approaching high overdispersion, male and female worms distributed together.)

$$y_t(m, 1) = m^2 / [(1 + m)(2 + m)] \quad [7]$$

where  $y$  is the fraction of worm pairs in the human population,  $m$  is the mean worm load per person, and  $k$  is the dispersion of clumping parameter.<sup>6</sup> We have found at this stage that as  $k$  approaches 0, it is difficult to solve the general equation and obtain meaningful results.

By substituting baseline prevalence  $y_t$  into equation and solving for  $m$ , we obtained mean worm loads for each age group in Sectors I through IV. Then, for each sector we graphed  $m$  versus geometric mean eggs to see if there were a functional relationship between these two variables (see Figure 9). We are still exploring what the appropriate functional relationship might be by graphing  $m$  against arithmetic mean egg count. When we establish that relationship, we believe it will be possible to use egg counts to estimate incidence. By multiplying the estimated egg counts by the number of positive persons in the age group,

---

<sup>6</sup>The concern about how the sexes of worms are distributed has been described in detail in May, and Bradley and May (42). Two extremes are postulated. In one, sex is essentially random and each worm has equal probability of being male or female. In this case, it is assumed that snails are heavily infected. The other extreme is that female and male worms are distributed each with a separate expression to describe its distribution. A condition for this case is that snails are lightly infected. It has been pointed out that the differences between the extremes are "non-trivial." We believe that we have a situation warranting use of the first case since transmission is fairly intense.

we obtain a value for the total number of eggs passed in the age group. The estimate of incidence can then be interpreted as a contamination/exposure rate. This value can be used in the transmission equation [2] as the infection coefficient multiplied by the number of uninfected persons  $(1 - y_t)$ . The new value of predicted prevalence can then be used to estimate mean worm load. The process could continue as described. The reasons for such additional complications would be to use egg count information while maintaining the recursive model. The egg count function would change as a result of controls and this could be easily reflected in the model.

Population movement or migration in and out of an area has been the subject of extensive economic and geographical studies (43). In epidemiological studies population movements have been shown to influence the spread of disease. Water resource development projects, especially man-made lakes in Africa, exemplify this problem. In Lake Volta, Ghana, low schistosomiasis prevalence levels were found before the lake was created (44). The creation of the lake led to the elimination of the downstream and delta fisheries. The infected fisherman moved to the lakeshore in large numbers (80,000) bringing their skills and schistosomiasis with them. In addition, the flooded lands created ideal snail habitats; snail populations rose rapidly. As a result of the increased snail population and infected migrants, schistosomiasis prevalence levels rose to 80 percent within five years of the creation of the lakes (45). The Gezira irrigation scheme in the Sudan is yet another example of this phenomenon. As the desert became arable with water channeled from the Nile, migrants from Egypt

moved to the area bringing schistosomes and snails with them. In 1925, when the scheme was completed, there was no schistosomiasis in the area. In three years, the population prevalence levels reached 30 percent (46).

Without attempting to understand the motives of schistosome-bearing migrants, it is still important to try to account for their influence in an area, that is, the introduction of infection to a previously uninfected area, or the increase of infection in an infected area. The average rate of movement of infected migrants into and from an area could be used to weight the population size values in the model. However, since the process of movement of uninfected persons may be thought of as random and taking place at discrete times, one may consider use of a Markov chain process (47). That is, there are four states of movement: in-movement of infected persons ( $a_1$ ), in-movement of uninfected persons ( $a_2$ ), out-movement of infected persons ( $a_3$ ), and out-movement of uninfected persons ( $a_4$ ). One may estimate transition probabilities ( $P_{ij}$ ) of moving from one state to another so that at any given time the population consists of infected migrants, uninfected migrants, infected residents, uninfected residents. The probabilities are usually expressed in a matrix (T):

$$T = \begin{matrix} & \begin{matrix} a_1 & a_2 & a_3 & a_4 \end{matrix} \\ \begin{matrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{matrix} & \begin{pmatrix} P_{11} & P_{12} & P_{13} & P_{14} \\ P_{21} & P_{22} & P_{23} & P_{24} \\ P_{31} & P_{32} & P_{32} & P_{34} \\ P_{41} & P_{42} & P_{43} & P_{44} \end{pmatrix} \end{matrix}$$

where the sum of the elements of each row = 1 (48). If we start in state  $i$ , what we want to estimate is the probability of moving to state  $j$ . We need to have an estimate of those probabilities in order to account for the different movements, for example infected migrants becoming uninfected, moving out or staying infected. With the Tanzania data, it may be possible to determine the usefulness of such an approach in weighting the population size.

A simpler approach to account for the likelihood of new infections in an uninfected area might be to add a random variable to the regression equation. Random number computer programs exist and yield normally distributed random variables with mean 0 and variance 1. It may be possible to add this random variable ( $R$ ) to each variable in the regression equation:

$$\text{(Incidence)} \quad I = (B_0 + R_0) \left[ H^{(B_1 + R_1)} \times P^{(B_2 + R_2)} \right]$$

$$\text{(Reversion)} \quad B = .2 + (R_3 \times \text{standard deviation of } B \text{ if estimated})$$

This may oversimplify the stochastic processes involved in transmission. Nonetheless, because of its simplicity, we will test it for its reliability in predicting the situation in Tanzania, where migration superimposes random conditions for transmission.

#### Future Work

Further modeling work will include modifying the population estimates to include migration information, examining which egg count distribution is most appropriate for use in the model, and then establishing which variables provide necessary and sufficient information for predictive

modeling so as to establish a priority listing of variables to include in predictive analyses. The next report will also include results from the cost-effectiveness analyses. The final report will pull together the modeling and economic analyses to present the work in a form useful for water resource project evaluation.

References

1. Rosenfield, P.L., R.A. Smith, and M.G. Wolman. 1977. Development and Verification of a Schistosomiasis Transmission Model. Am J. Trop. Med. Hyg. 26 (3): 505-516; Rosenfield, P.L. 1975. Schistosomiasis Transmission Model (Washington Agency for International Development), 162 pp.
2. Rosenfield, P.L. 1977. "Water Project Effects," Contract with Agency for International Development, Project No. 931-1133, p. 2.
3. Lanoix, J.N. 1958. Relation Between Irrigation Engineering and Bilharziasis. Bull. WHO, 18:1011-1035.
4. Olivier, L.J. 1973. Evaluation of Control Programmes, in N. Ansari, ed., Epidemiology and Control of schistosomiasis (Baltimore, University Park Press), p. 609.
5. Ibid., p. 614.
6. Warren, K.S. 1973. Regulation of the prevalence and intensity of schistosomiasis in man: immunology or ecology? J. Inf. Diseases, 127: 595-609.
7. Ibid.
8. See series by Wilkins, H.A., et. al., 1977. *Schistosoma haematobium* in a Gambian community, Ann. Trop Med. Parasit., 71(1): 53-195.
9. Ritchie, L. 1973. Chemical control of snails, in N. Ansari, ed., Epidemiology and Control of Schistosomiasis (Baltimore, University Park Press), 458-532.
10. Observed in many areas studied by the Principal Investigator. See, for example, Chu, K.Y., J. Massoud, and F. Arfaa. 1968. Distribution and ecology of Bulinus Truncatus in Khuzestan, Iran. Bull. WHO 39(4): 607-632.
11. Muench, H. 1959. Catalytic Models in Epidemiology (Cambridge, Harvard University Press), 110 pp.
12. Rosenfield, et. al., op. cit., pp. 506-507.
13. Jordan, P. 1977. Schistosomiasis research to control. Am. J. Trop. Med. Hyg. 26: 877-886; Shiff, C.J. 1972. The value of incidence as a parameter for assessing control of Bilharziasis in Rhodesia. SCHISTO/WP/72.25, World Health Org., Geneva (unpubl.).
14. Wright, W. 1973. Geographical distribution of schistomes and their intermediate hosts, in N. Ansari, ed., Epidemiology and Control of Schistosomiasis (Baltimore University Park Press), p. 115.

15. Webbe, G. 1972. Control of Schistosomiasis in Ethiopia, Sudan, and East and West African Countries, in M. Miller, ed., Schistosomiasis (New Orleans, Tulane University), pp. 118-119.
16. Wright, op. cit., p. 116.
17. Ibid.
18. Ibid., p. 117.
19. Jordan, P. 1966. Medical Aspects in G. Webbe and P. Jordan, Recent Advances in Knowledge of Schistosomiasis in East Africa, Trans. Roy. Soc. Trop. Med. Hyg., 60 (2): 294-303.
20. Blue, R.N. and J.H. Weaver. 1977. A critical assessment of the Tanzanian model of development, Paper No. 30, Agricultural Development Council, p. 10.
21. Ibid., p. 13.
22. McCullough, F. and V.M. Eyakuze. 1973. WHO/Tanzania Schistosomiasis Pilot Control and Training Project, Mwanza District, Tanzania, Final Report. AFR/SCHIST/29, World Health Organization, Geneva (unpubl.). p. iii.
23. Ibid., p. iii.
24. Ibid., p. iii.
25. Ibid., p. iii.
26. Ibid., p. iii
27. Ibid., p. 1.
28. Ibid., p. 1.
29. Ibid., p. iii.
30. Ibid., pp. 2,5.
31. Ibid., p. 5.
32. Ibid., pp. 5,6.
33. Dalton, P. 1976. A Socioecological Approach to the Control of Schistosoma mansoni in St. Lucia. Bull. WHO 54: 587-595.
34. Malcolm, D.W. 1953. Sukumaland: An African People and their Country (London, Oxford University Press), 224 pp.

35. Ruysenaars, J., G. Van Etten, and F. McCullough. 1973. Population Movement in relation to the spread and control of schistosomiasis in Sukumaland, Tanzania. Trop. Geog. Med., 25: 179-186.
36. May, R. 1977. Togetherness among schistosomes, Math. Bios. (in press); Bradley, D.J. and R.M. May, 1977. Consequences of helminth aggregation for the dynamics of schistosomiasis, Trans. Roy. Soc. Trop. Med. Hyg. (in press).
37. MacDonald, G. 1965. The dynamics of helminth infections, with special reference to schistosomes, Trans. Roy. Soc. Trop. Med. Hyg., 59: 489-506.
38. May, Bradley and May, op. cit.
39. Bradley and May, op. cit., p. 23.
40. Ibid.
41. Ibid.
42. Ibid.
43. Greenwood, M.J. 1975. Research on internal migration in the United States: a survey. J. Econ. Lit., 13: 397-433.
44. Annex 15 from UNDP/WHO Project Review Mission, 28 November-11 December, 1975, Accra.
45. Ibid.
46. Humphreys, R.M. 1932. Vesical schistosomiasis in the Gezira irrigated area of the Sudan, Sept. 1925 - Dec. 1974. Trans. Roy. Soc. Trop. Med. Hyg., 26: 241-252.
47. Kemeny, J.G., J.L. Snell, G.L. Thompson. 1957. Introduction to Finite Mathematics (Englewood Cliffs, Prentice-Hall), pp. 171-177.
48. Ibid., p. 172.

**FIGURES AND TABLES**

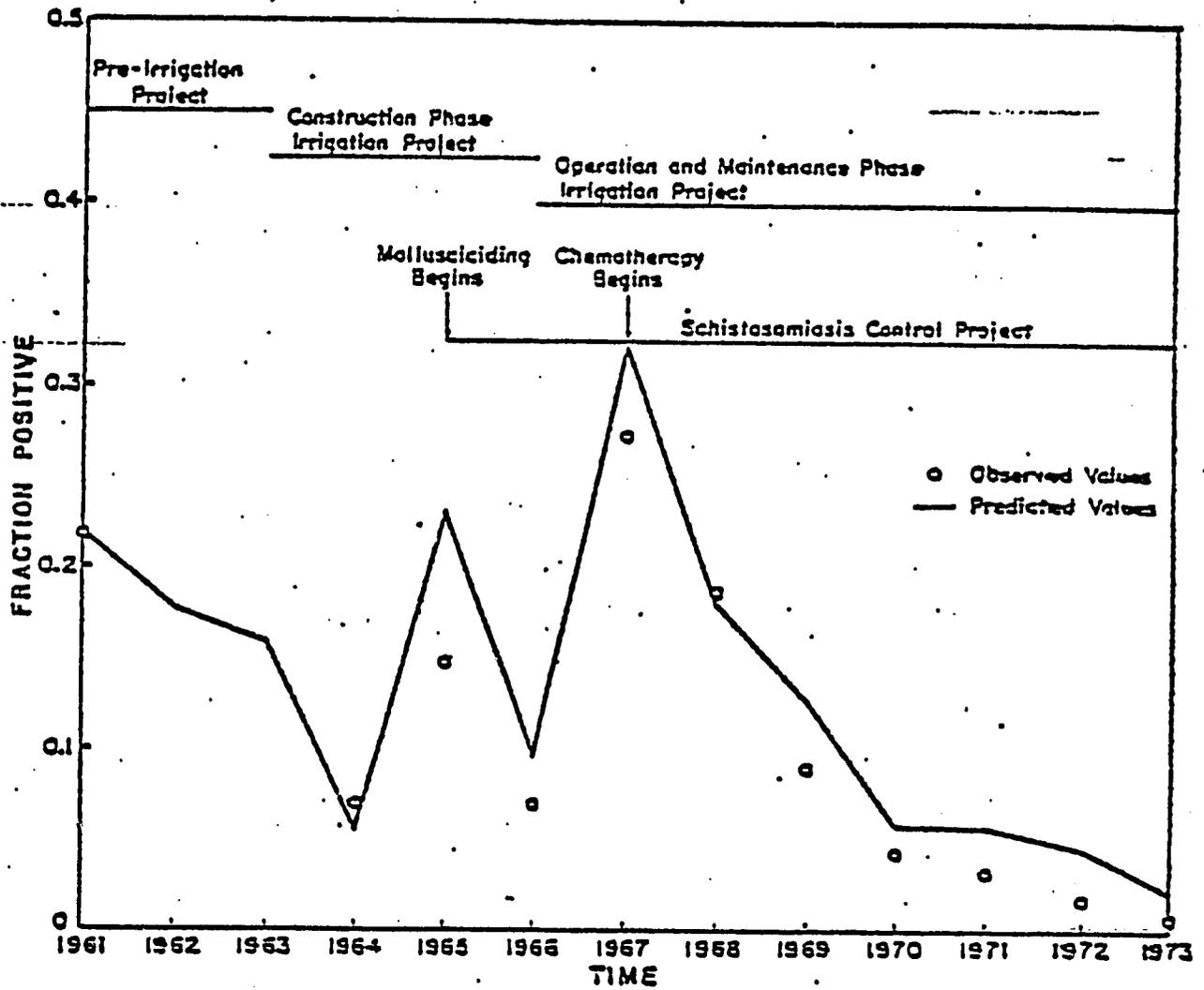


Figure 1A.

Testing Results of Schistosomiasis Model with Iranian Data. Predicted values of prevalence (connected line) are compared with observed values (open circles) for years of irrigation project construction and operation, and Bilharziasis Control Project operations.

RICHE FOND ALL AGES COMBINED

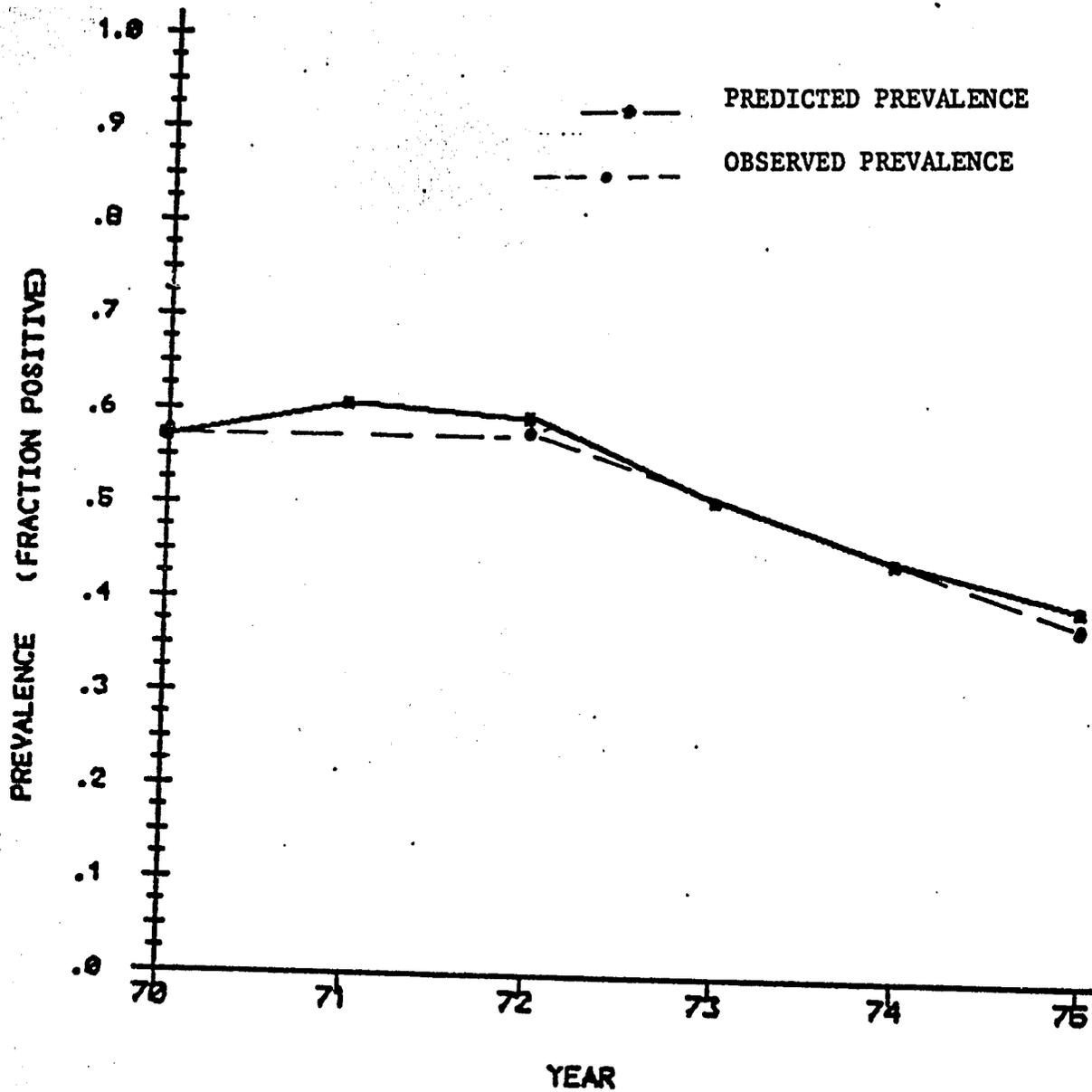
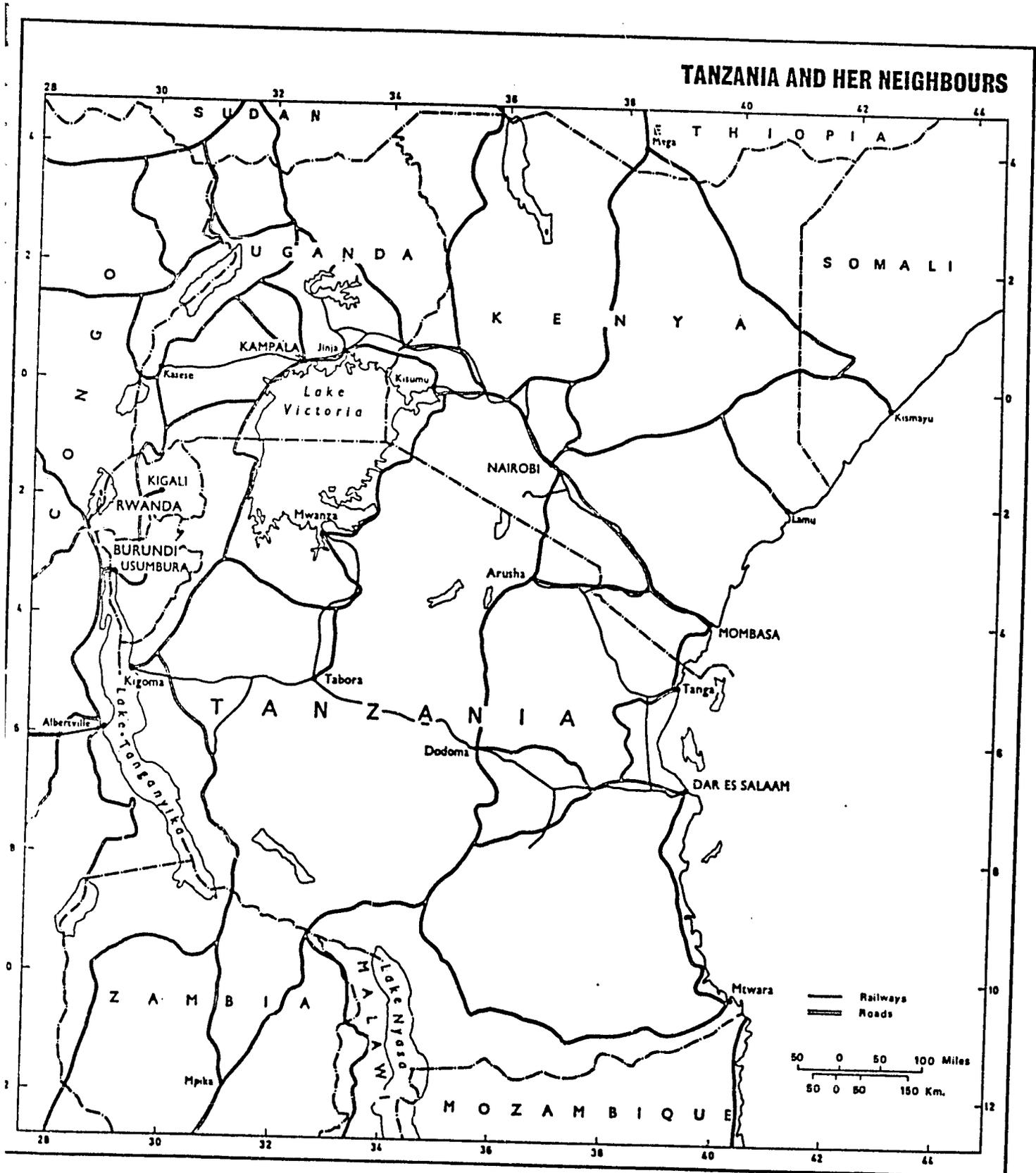


Figure 1B.

Results of Transmission Model Testing with Riche Fond, St. Lucia Data. Predicted prevalence values are compared with observed values. Water supply installation was started in 1970 and completed by 1973.

Figure 2. Map of Tanzania. From Berry, L., ed. 1971. Tanzania in Maps (London, University of London Press). p.13.



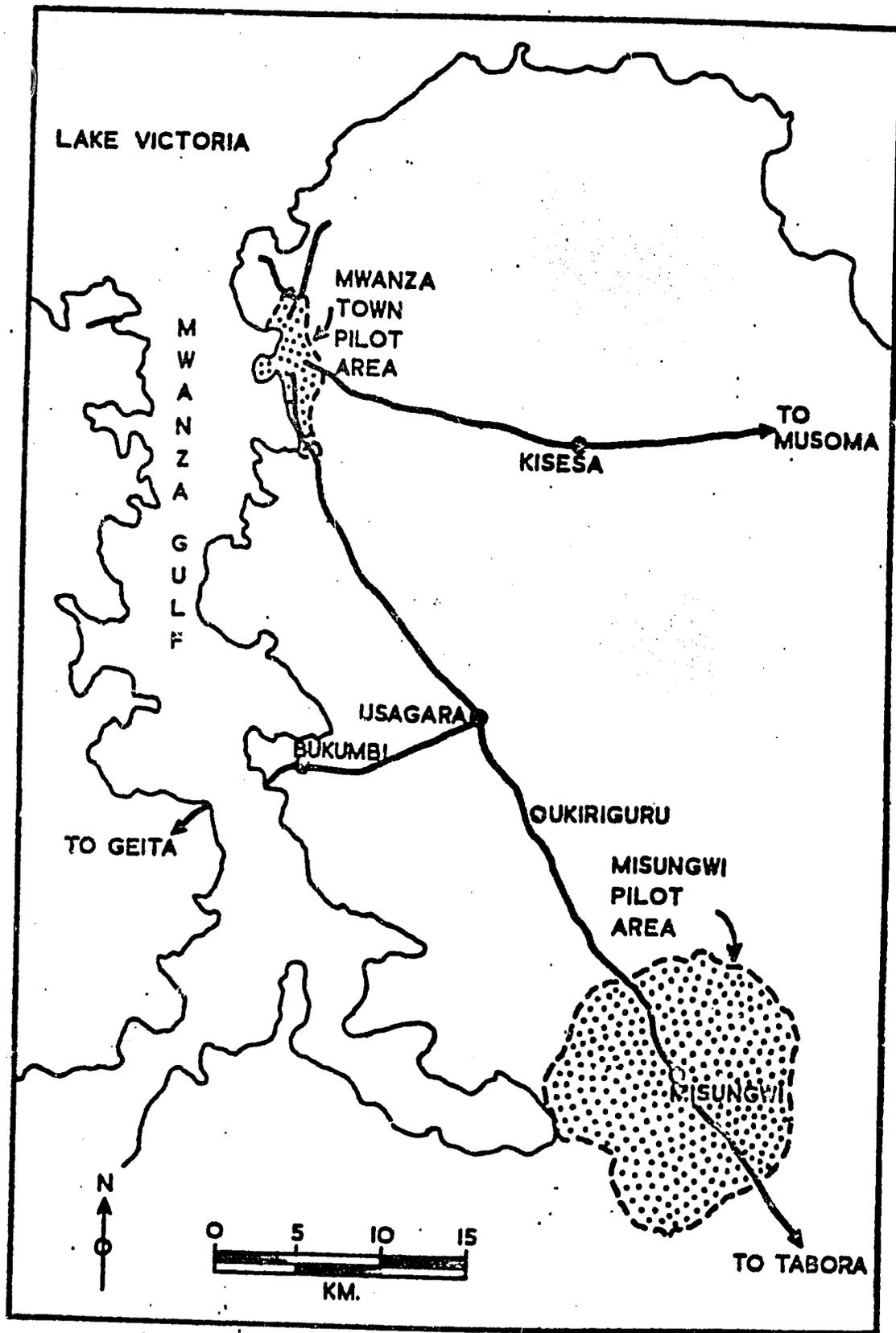


Figure 3. Showing the location of the Misungwi and Mwanza pilot areas.

From: McCullough, F. and V. Eyakuze, 1973. WHO/Tanzania Schistosomiasis Pilot Control and Training Project, Mwanza District, Tanzania, Final Report, AFR/SCHIST/29, WHO, Geneva.

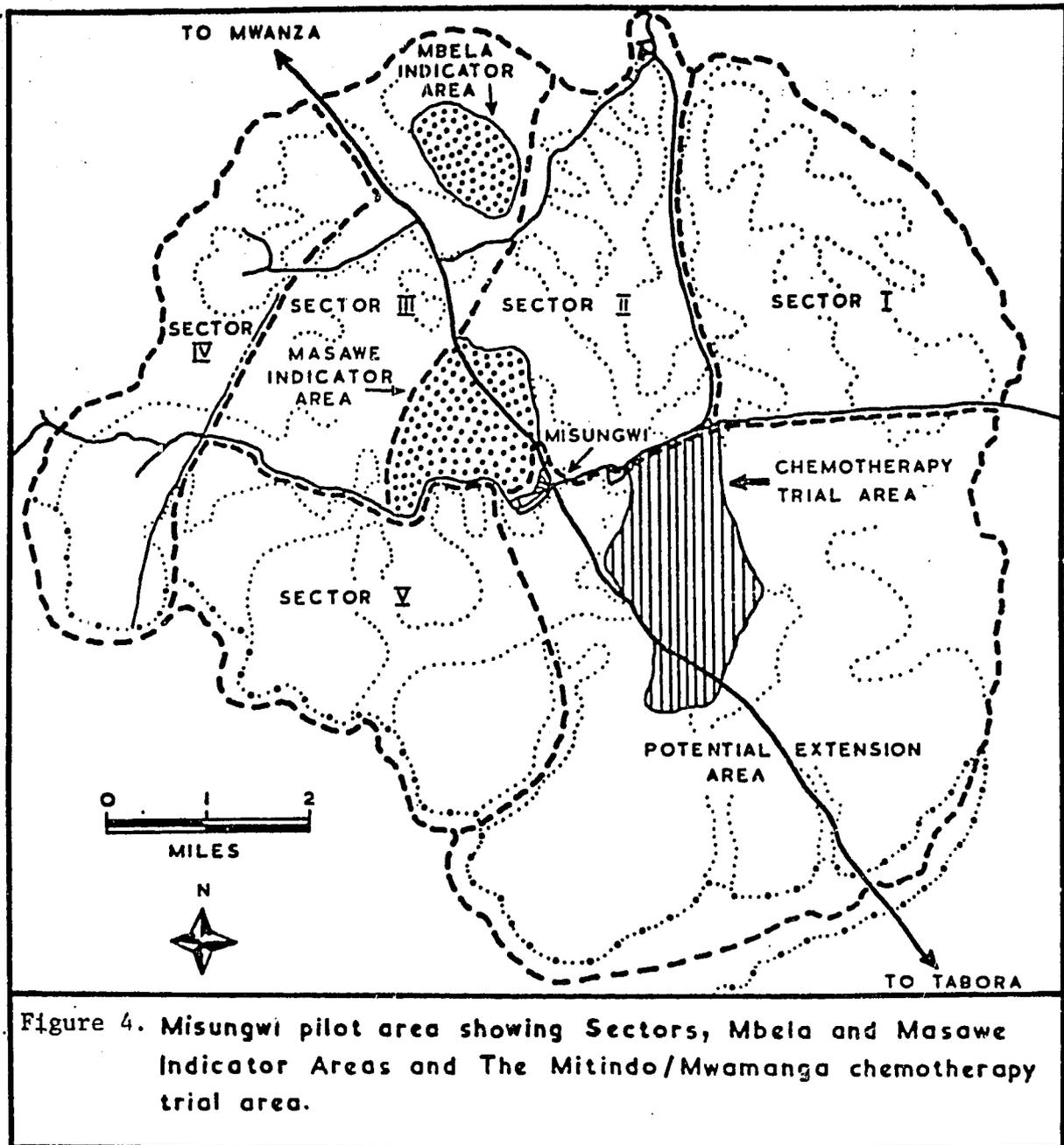


Figure 4. Misungwi pilot area showing Sectors, Mbela and Masawe Indicator Areas and The Mitindo/Mwamanga chemotherapy trial area.

From: McCullough, F. and V. Eyakuze, 1973. WHO/Tanzania Schistosomiasis Pilot Control and Training Project, Mwanza District, Tanzania, Final Report, AFR/SCHIST/29, WHO, Geneva.

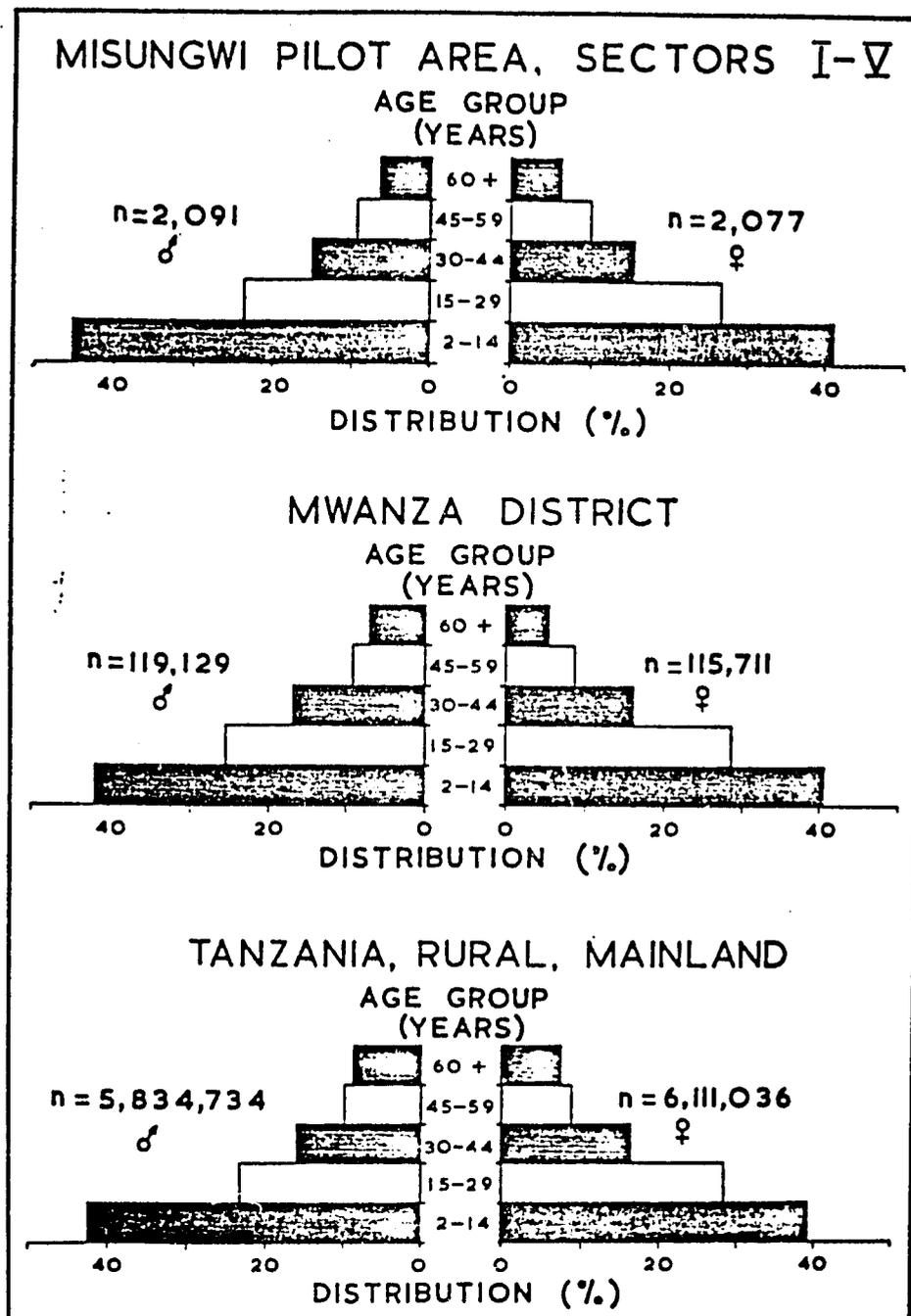
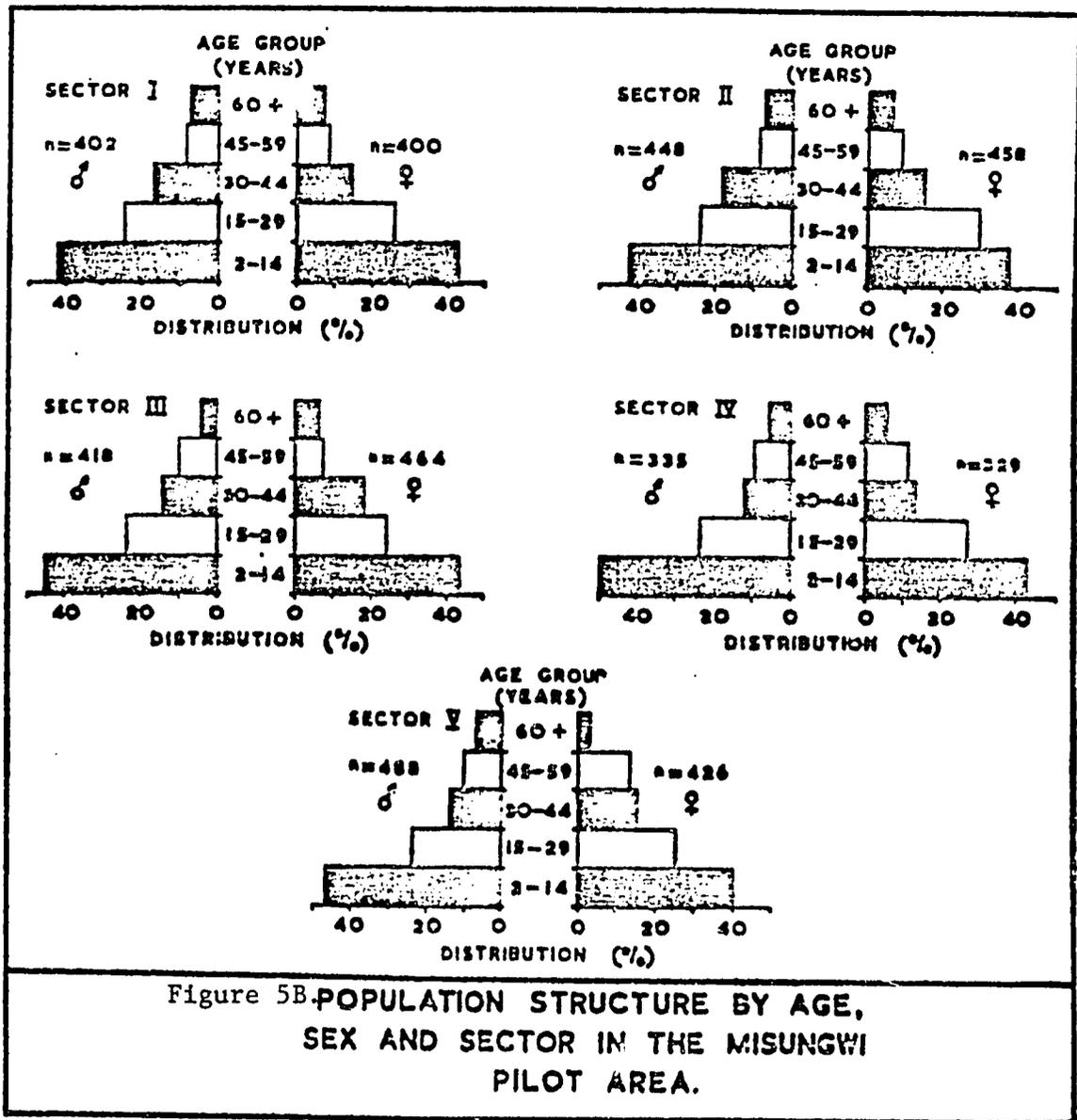


Figure 5A Comparison of the age/sex structure of the populations in the Misungwi pilot area, Mwanza District and mainland Tanzania.

From: McCullough, F. and V. Eyakuze, 1973. WHO/Tanzania Schistosomiasis Pilot Control and Training Project, Mwanza District, Tanzania, Final Report, AFR/SCHIST/29, WHO, Geneva.



From: McCullough, F. and V. Eyakuze, 1973. WHO/Tanzania Schistosomiasis Pilot Control and Training Project, Mwanza District, Tanzania, Final Report, AFR/SCHIST/29, WHO, Geneva.

**KEY**

- HOUSEHOLD
- HABITAT (X=dam)
- ..... edge of valley
- paths
- main road

(map reduced 35%)

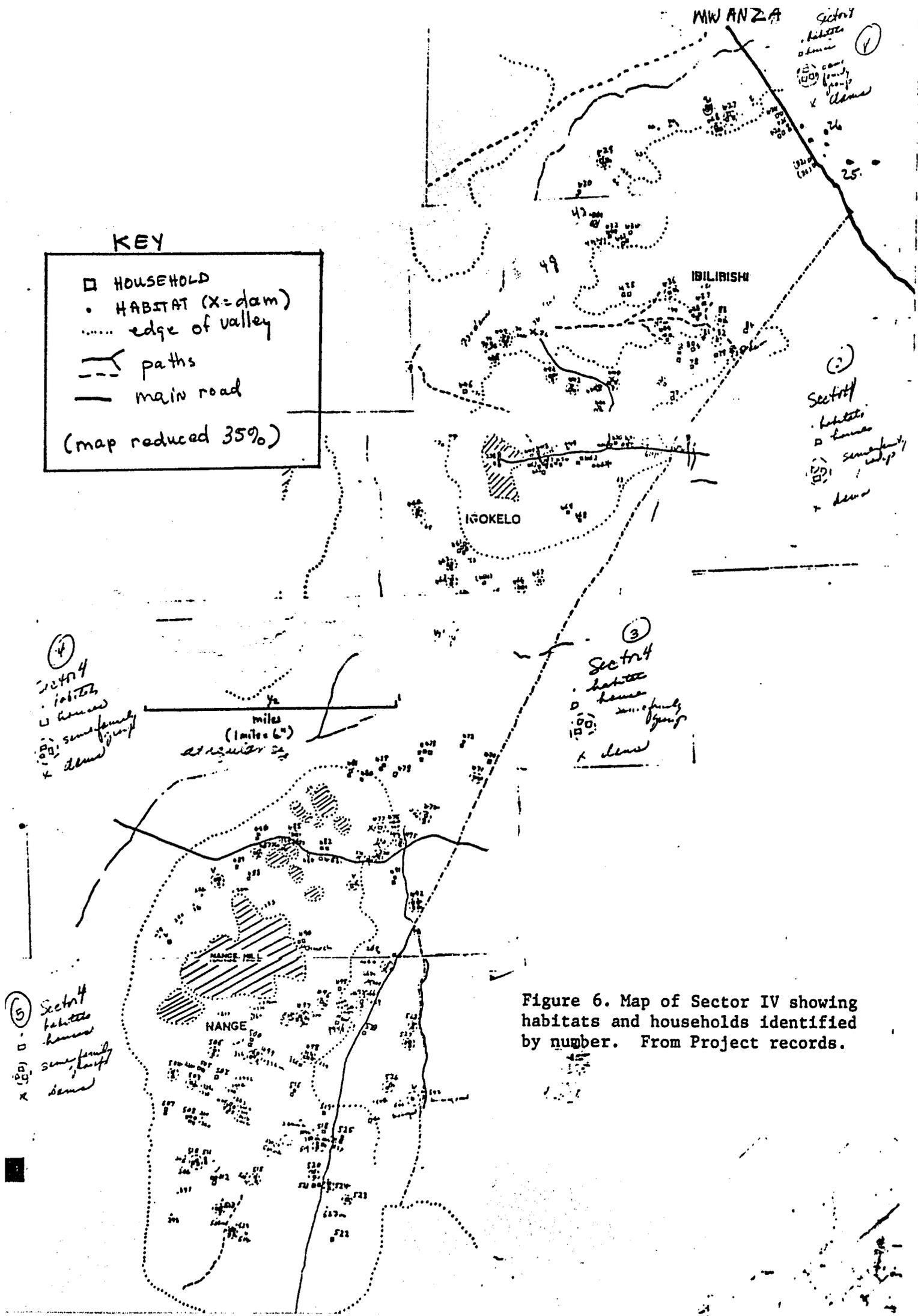


Figure 6. Map of Sector IV showing habitats and households identified by number. From Project records.

Table 1. Results from the different regressions are presented in this table. The regressions were run for distances of 1/2 and 1/8 mile (assumptions of habitat accessibility) for different age groupings. The variables in the regressions were: I, incidence; V, volume of accessible habitats (cu.ft.); P, number of infected persons; E, arithmetic mean egg count; GE, geometric mean egg count; PM, perimeter of accessible habitats (ft.); i, age group;  $\beta_0$ ,  $\beta_1$ ,  $\beta_2$ , the regression coefficients and exponents specific for each equation.  $R^2$  is the percentage of variance in the dependent variable (I) explained by the equation; F is the F-statistic which indicates the significance of the  $R^2$ , the numbers in parentheses (2,5) are the degrees of freedom; "t" is the statistic which indicates significance of the  $\beta$ 's at 5 degrees of freedom (which may be somewhat unreliable at small sample sizes). Table 1 A. Results for 1/8 mile runs. Table 2 B. Results for 1/2 mile runs.

Age Groupings	$I_i = \beta_0 (V_i^{\beta_1} \times P_i^{\beta_2})$ t= t= t=	$I_i = \beta_0 (V_i^{\beta_1} \times E_i^{\beta_2})$ t= t= t=	$I_i = \beta_0 (V_i^{\beta_1} \times GE_i^{\beta_2})$ t= t= t=	$I_i = \beta_0 (PM_i^{\beta_1} \times P_i^{\beta_2})$ t= t= t=	$I_i = \beta_0 (PM_i^{\beta_1} \times E_i^{\beta_2})$ t= t= t=	$I_i = \beta_0 (PM_i^{\beta_1} \times GE_i^{\beta_2})$ t= t= t=
	$R^2 =$	$R^2 =$	$R^2 =$	$R^2 =$	$R^2 =$	$R^2 =$
2,3,4,5, 6,7,8,9	$I = .261 (V_i^{-.09} \times P_i^{.513})$ -.45 1.58 -.29 $R^2 = .355$ $F(2,5) = 1.37$	$I = .003 (V_i^{.1} \times E_i^{.75})$ -1.70 1.72 .36 $R^2 = .392$ $F(2,5) = 1.61$	$I = .812 (V_i^{-.17} \times GE_i^{.45})$ -.069 1.81 -.533 $R^2 = .415$ $F(2,5) = 1.78$			
2,3-4, 5-8,9	$I = .054 (V_i^{.07} \times P_i^{.49})$ -2.19 3.38* .50 $R^2 = .772$ $F(2,5) = 8.473^*$	$I = .001 (V_i^{.25} \times E_i^{.71})$ -5.16 4.00* 2.27 $R^2 = .822$ $F(2,5) = 11.515^*$	$I = .164 (V_i^{-.55} \times GE_i^{.43})$ -1.77 5.18** -.05 $R^2 = .882$ $F(2,5) = 18.736^{**}$	$I = .009 (PM_i^{.32} \times P_i^{.47})$ -2.27 3.84* 1.23 $R^2 = .816$ $F(2,5) = 11.106^*$	$I = .003 (PM_i^{.48} \times E_i^{.57})$ -3.65* 3.29* 1.71 $R^2 = .771$ $F(2,5) = 8.434^*$	$I = .114 (PM_i^{-.04} \times GE_i^{.42})$ -1.253 5.097** 1.79 $R^2 = .883$ $F(2,5) = 18.860^{**}$
0-4, 5-9	$I = .166 (V_i^{-.02} \times P_i^{.42})$ -1.11 -.13 2.41 $R^2 = .585$ $F(2,5) = 3.52$	$I = .006 (V_i^{.13} \times E_i^{.55})$ -2.56* .85 2.12 $R^2 = .528$ $F(2,5) = 2.798$	$I = .406 (V_i^{-.08} \times GE_i^{.30})$ -.573 -.492 2.82* $R^2 = .65$ $F(2,5) = 4.71$			

\* Significant at  $P < .05$   
 \*\* Significant at  $P < .01$

TABLE 1 B.

2,3-4,  
5-8,9

$$I = 3.09 \times 10^{-6} (P^{.38} \times PM^{.84})$$

$$t = -1.82 \quad t = 2.58 \quad t = 1.42$$

$$R^2 = .830$$

$$F(2, 5) = 12.187$$

$$I = 1.6 \times 10^{-6} (P^{.36} \times V^{.91})$$

$$t = -2.23 \quad t = 2.67 \quad t = 1.85$$

$$R^2 = .858$$

$$F(2, 5) = 15.133$$

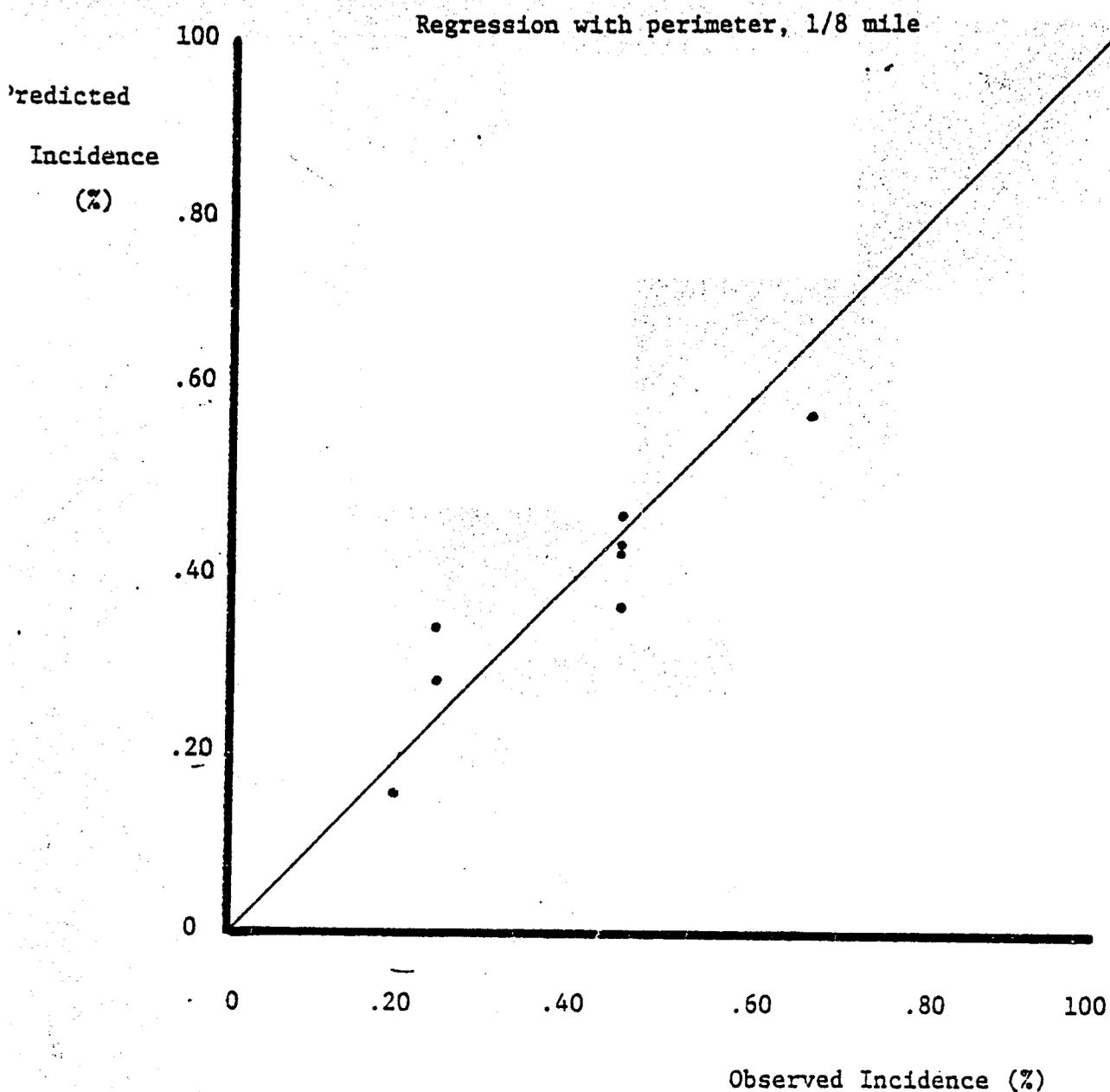
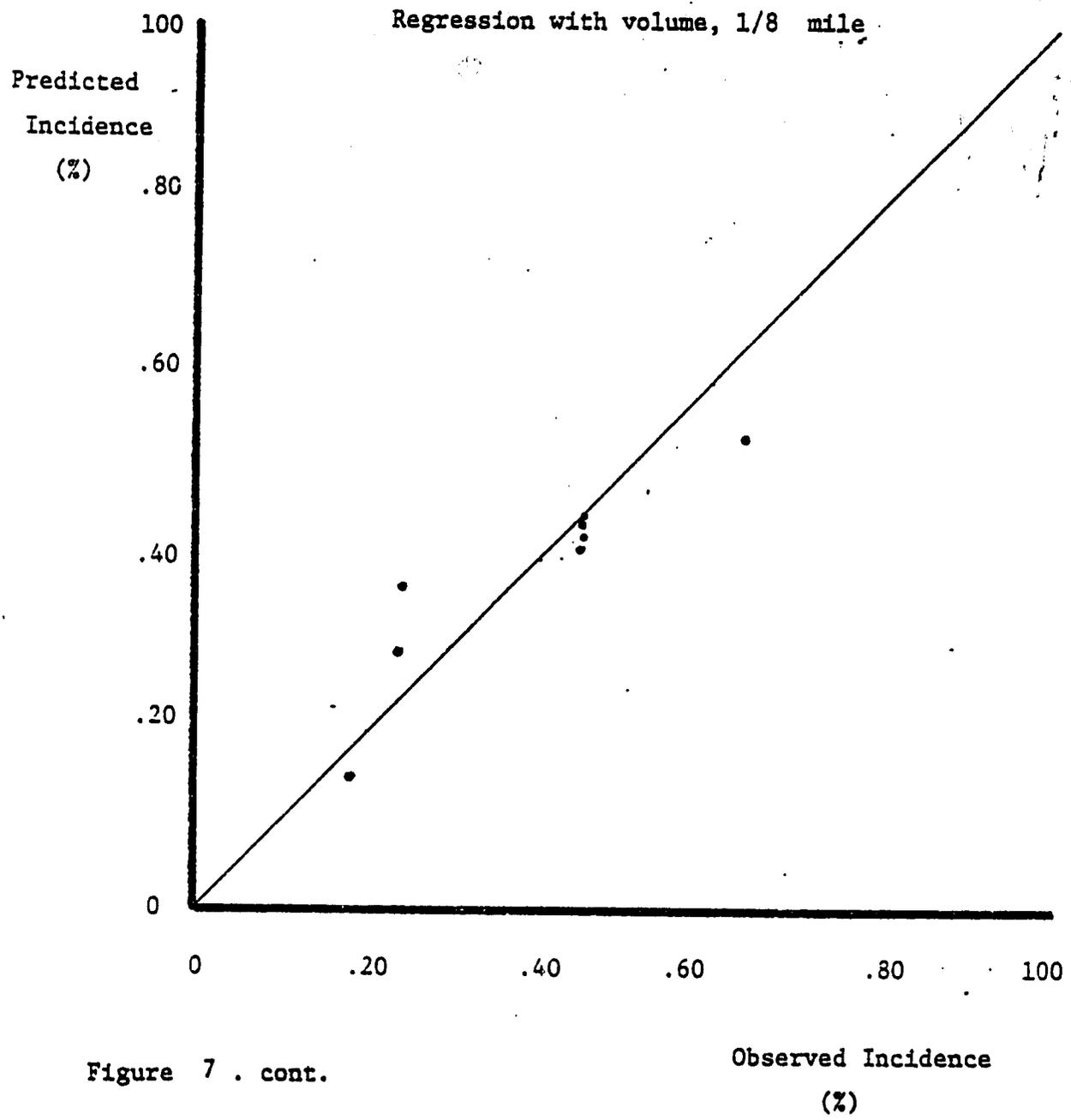


Figure 7 . Results of regression analyses for perimeter and volume under 1/8 and 1/2 mile accessibility assumptions. On the graph, predicted incidence values are plotted against observed values to show the scatter from the 45 line ( indication of closeness of fit of prediction: the less scatter from the line, the closer the fit).



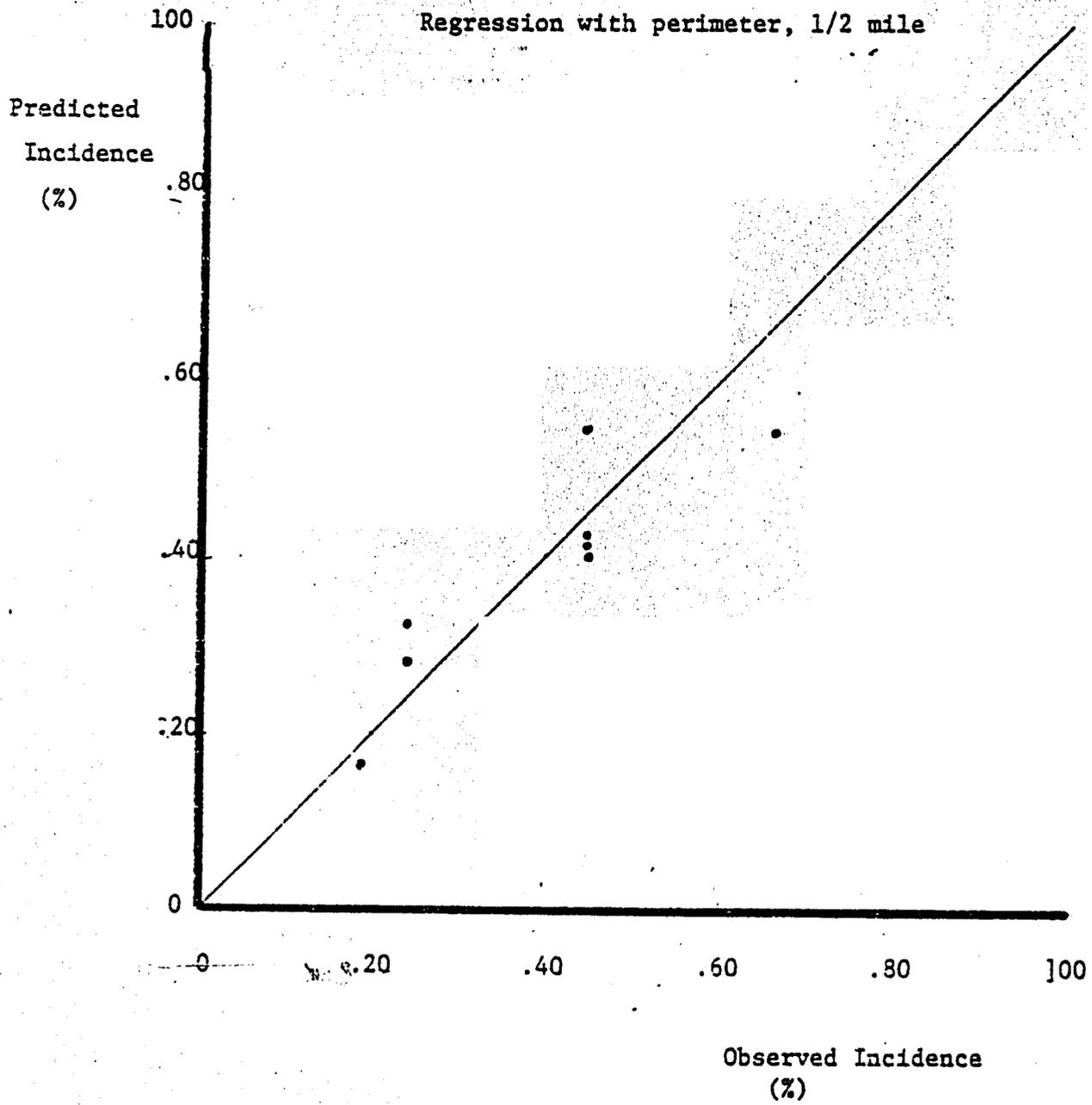


Figure 7 . cont.

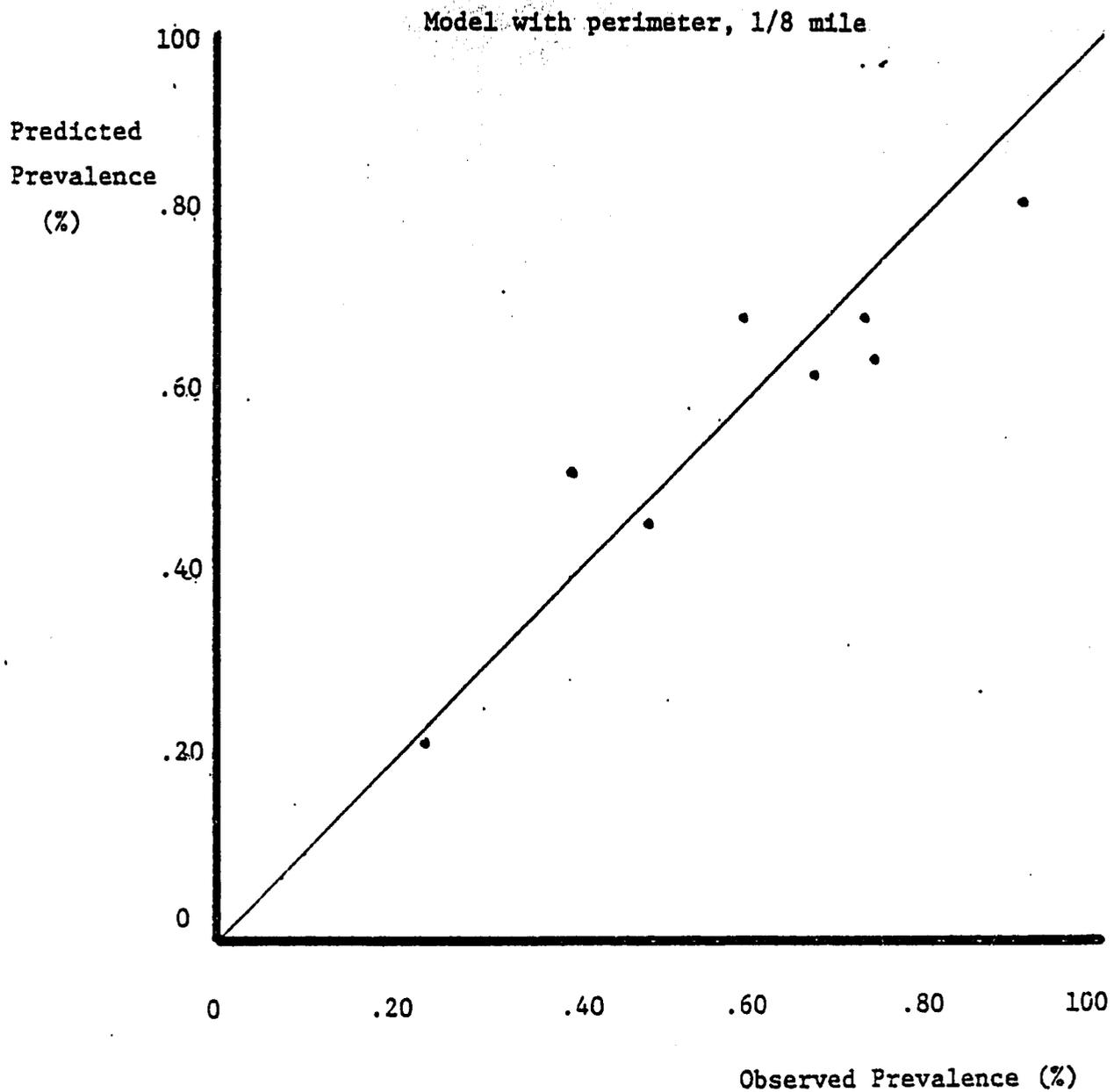


Figure 8 . Results of model runs for perimeter and volume under 1/8 and 1/2 mile accessibility assumptions. On the graph, predicted prevalence values are plotted against observed values to show the scatter from the 45 line ( indication of closeness of fit of prediction: the less scatter, the closer the fit).

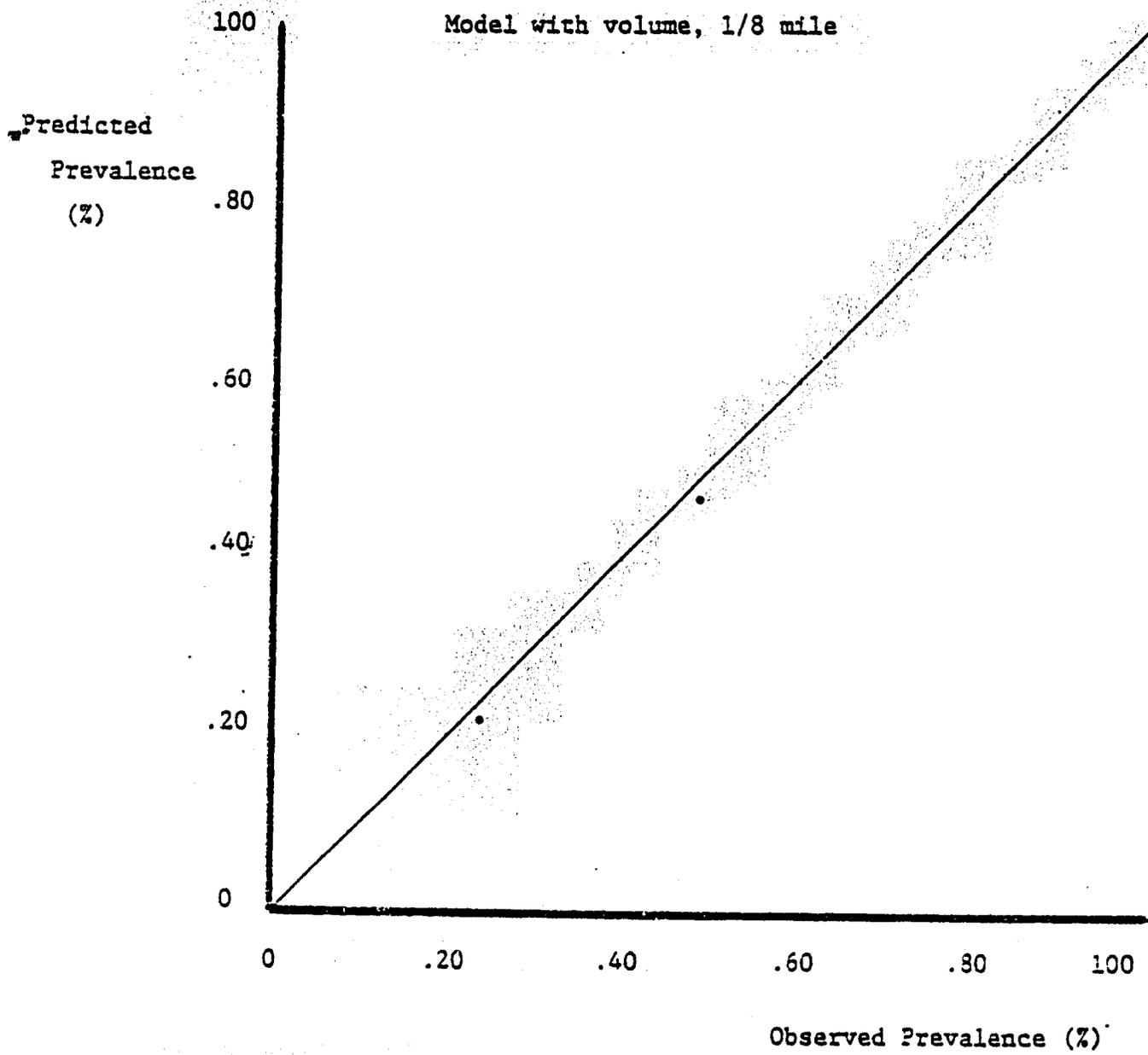


Figure 10

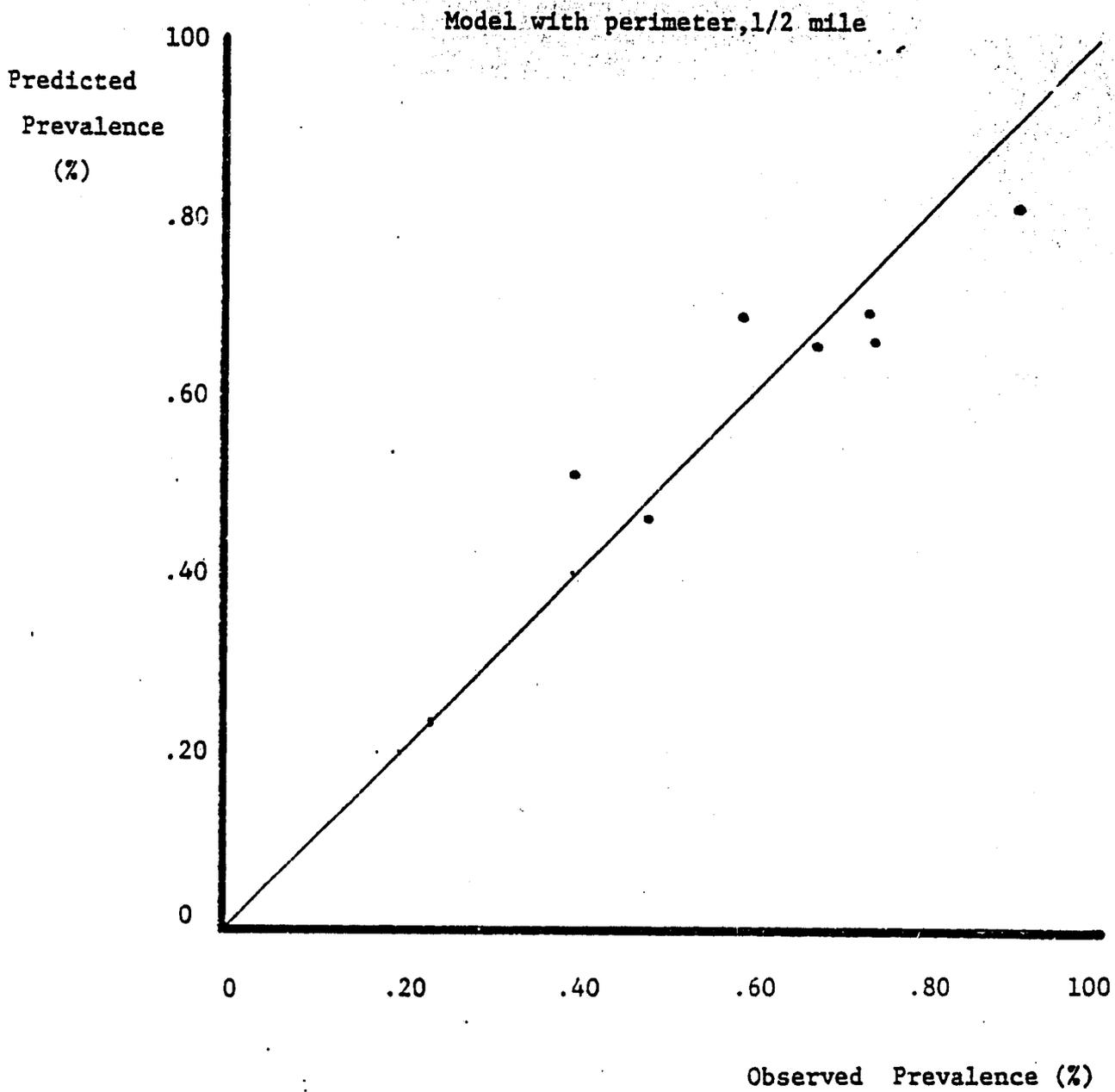


Figure 8 . cont.

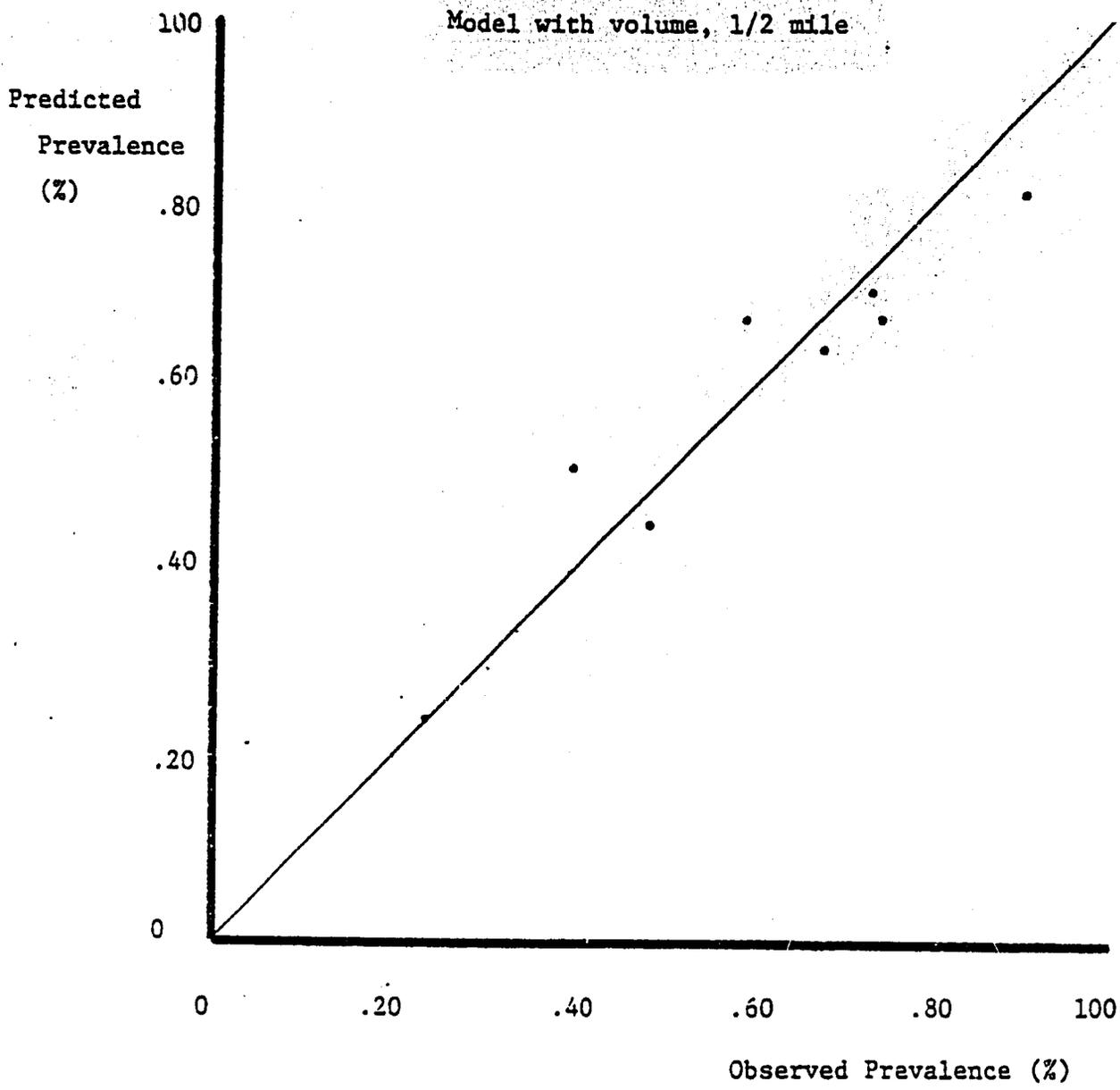
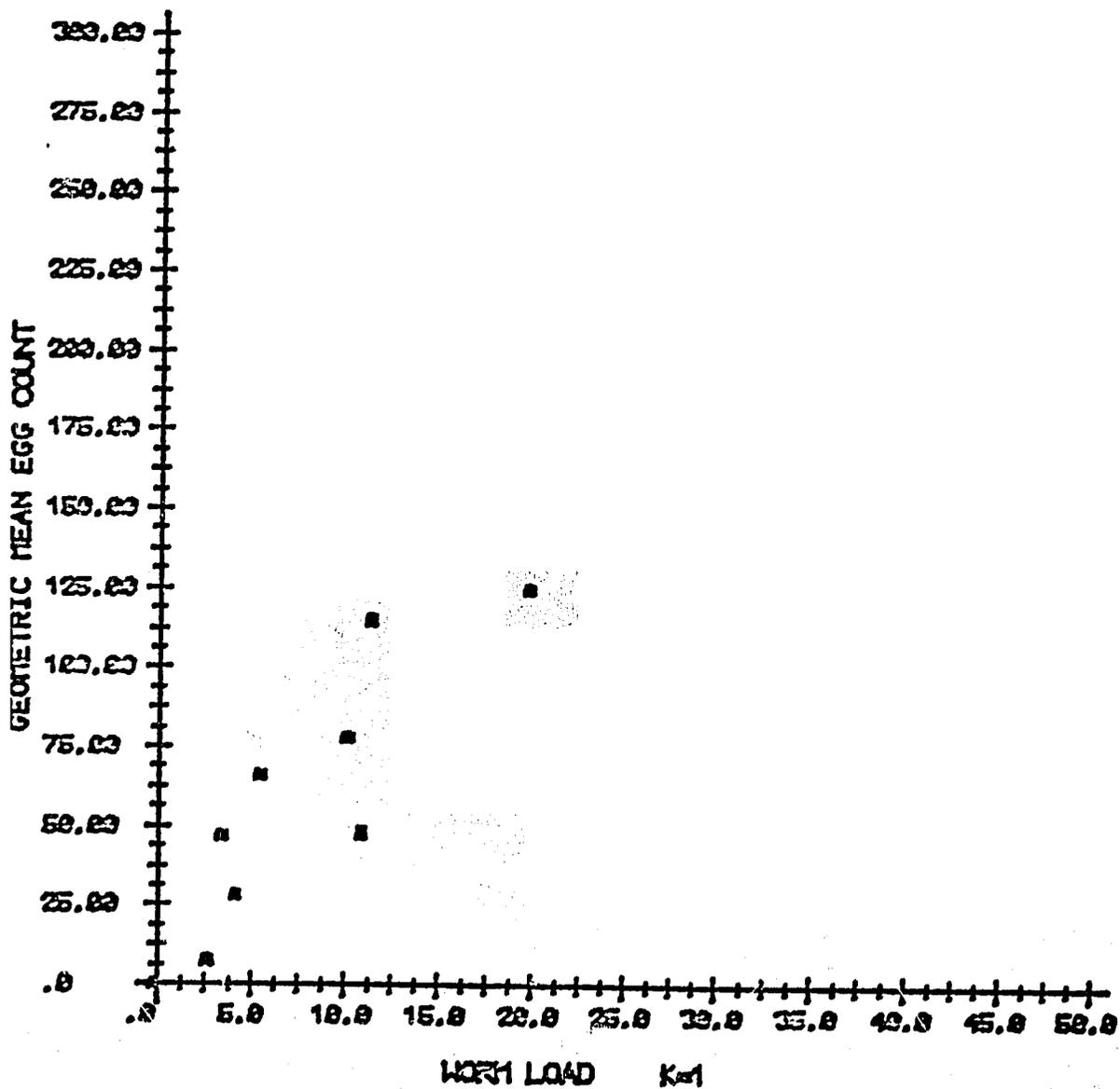


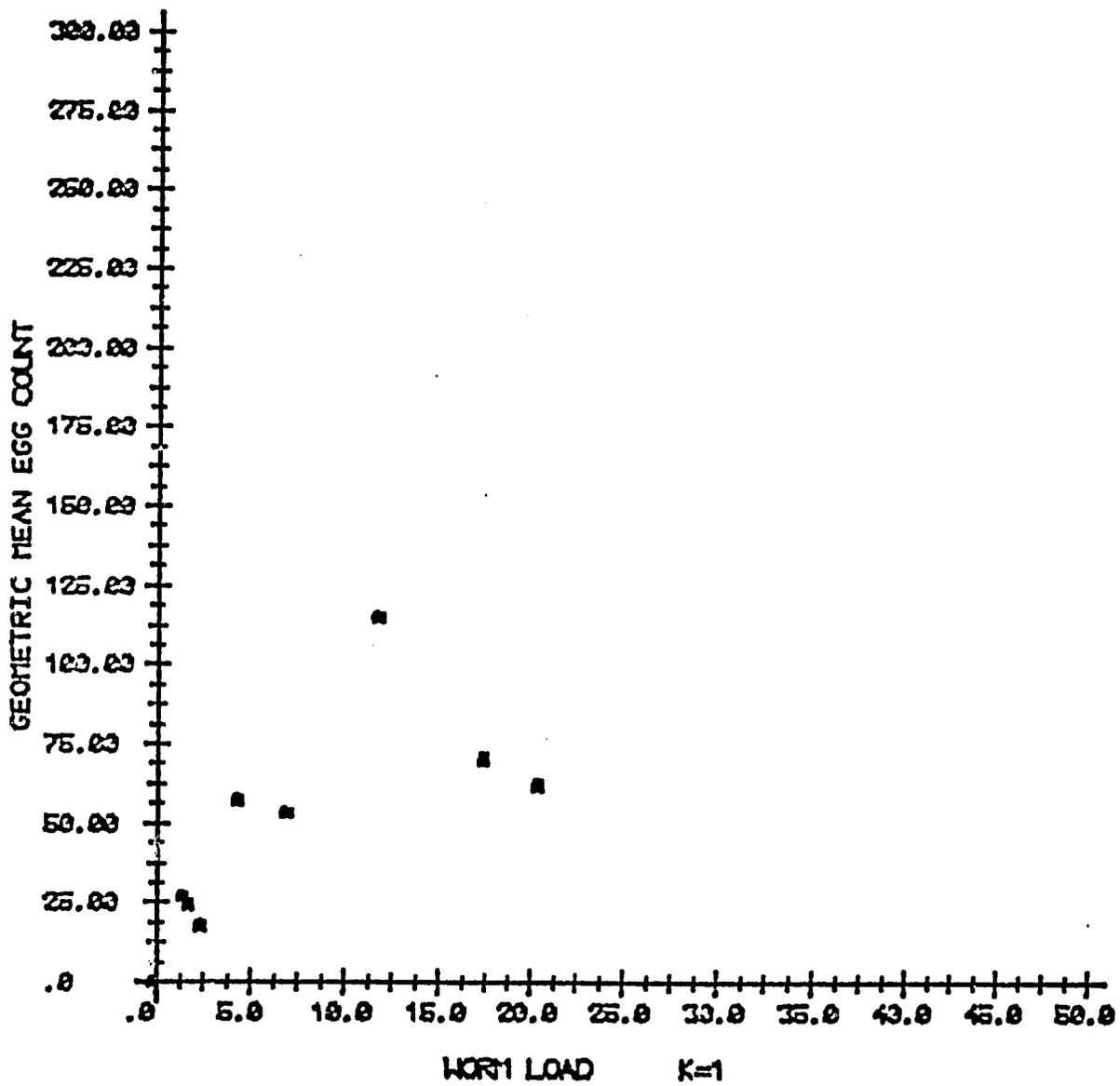
Figure 8 . cont.

Figure 9 (23 pages). These plots demonstrate the relationship between geometric mean egg count and mean worm load for the different sectors on a sex specific basis in Tanzania 001. There are two sets of plots. The first one shows results of calculating mean worm load from prevalence values under the assumption of a worm distribution approaching a negative binomial ( $k=1$ ). The second set shows results under the assumption of a Poisson distribution:  $k$  approaches infinity ( $k \rightarrow \infty$ ). The results are not fitted by a line but do show similar tendencies over the sectors and sexes for each distribution. Where  $k=1$ , the line of best fit would go through the origin, indicating that at 0 worms there are 0 eggs passed. Where  $k \rightarrow \infty$ , the line of best fit would pass through the x-axis, indicating that there may be some worms in humans before eggs are passed. That result is of interest since the worm load in this instance was estimated by prevalence, a measure determined by the presence or absence of eggs. These results require further study.

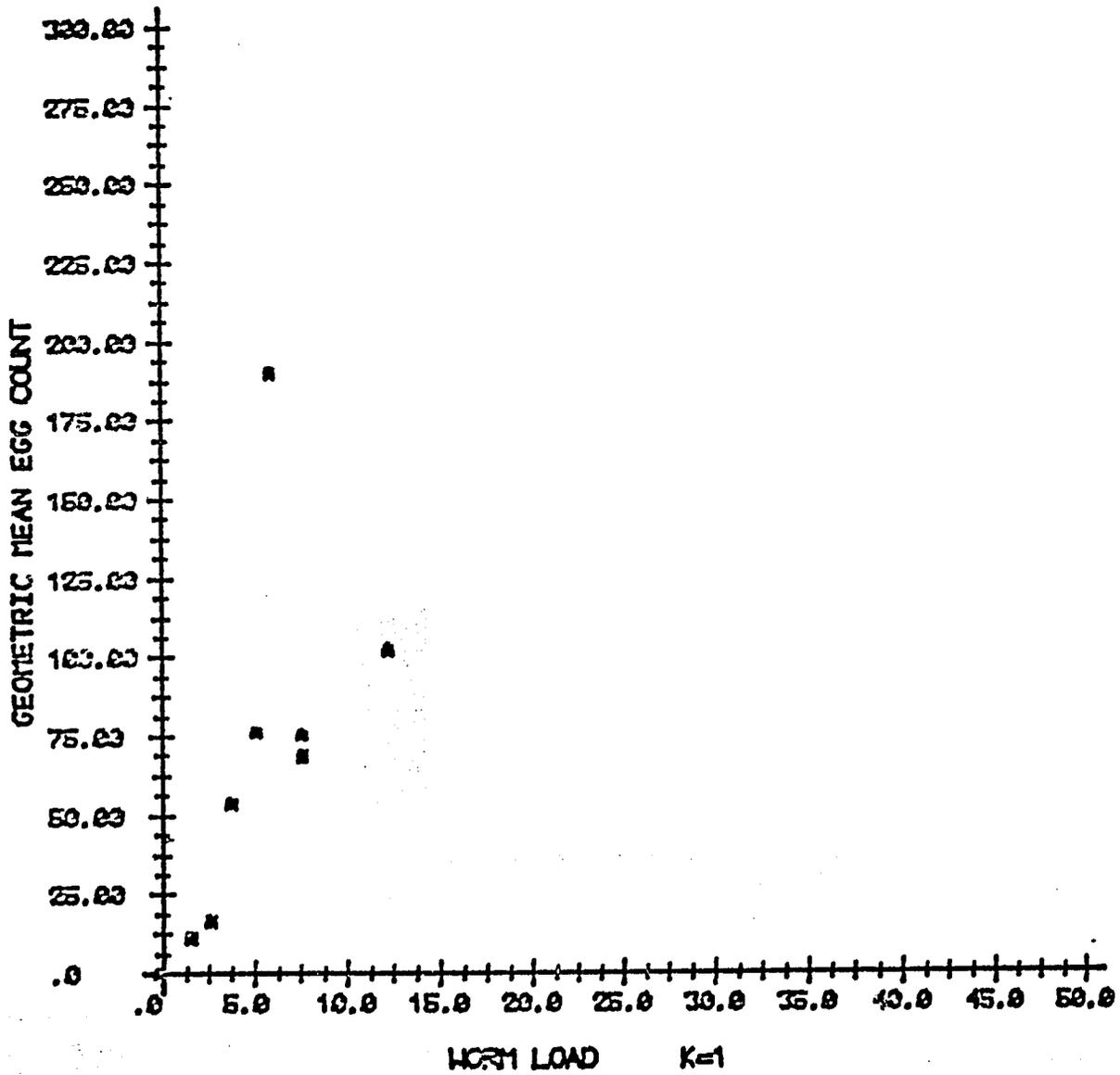
SECTOR 1 MALE AND FEMALE COMBINED



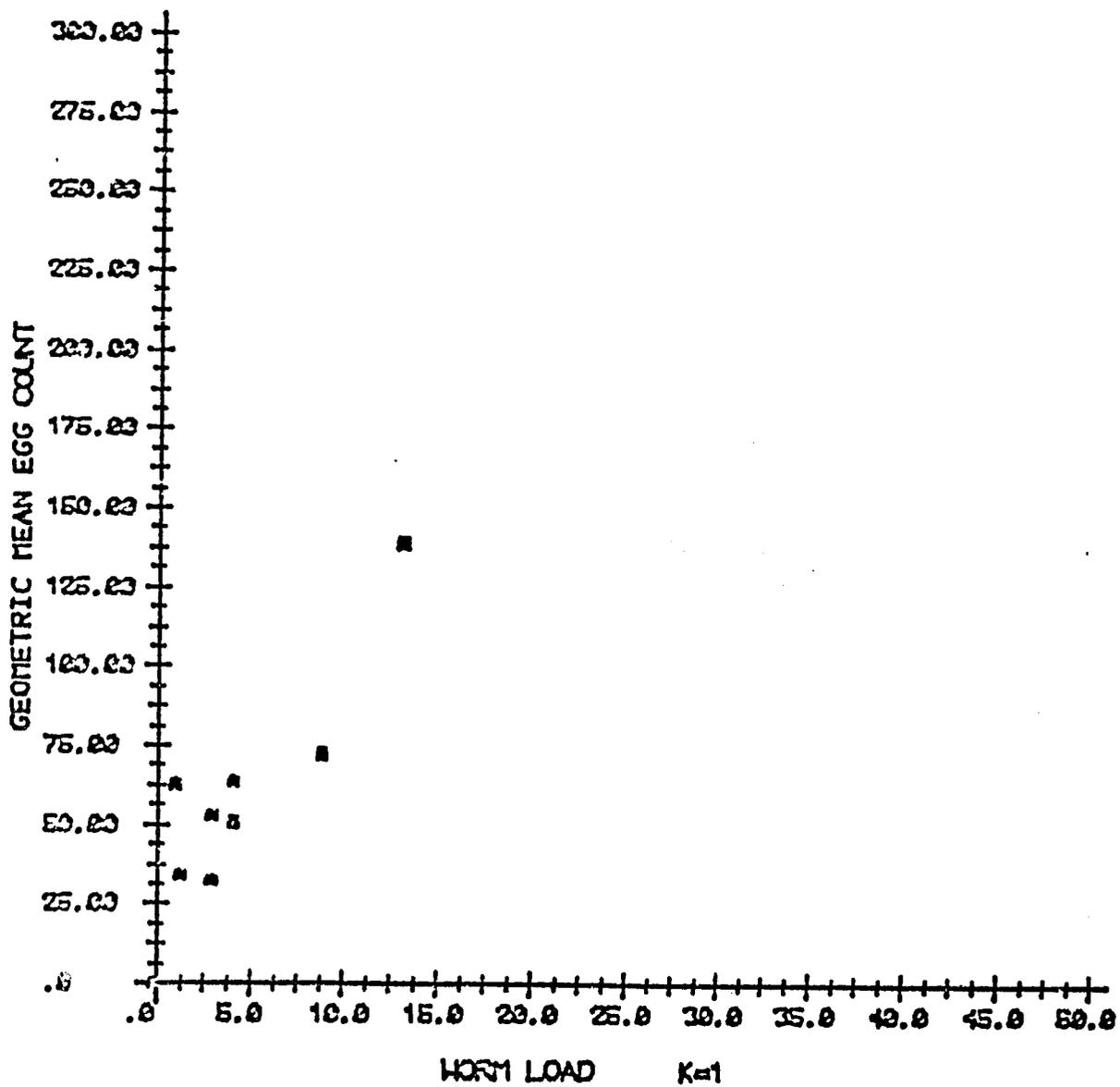
SECTOR 2      MALES AND FEMALES COMBINED



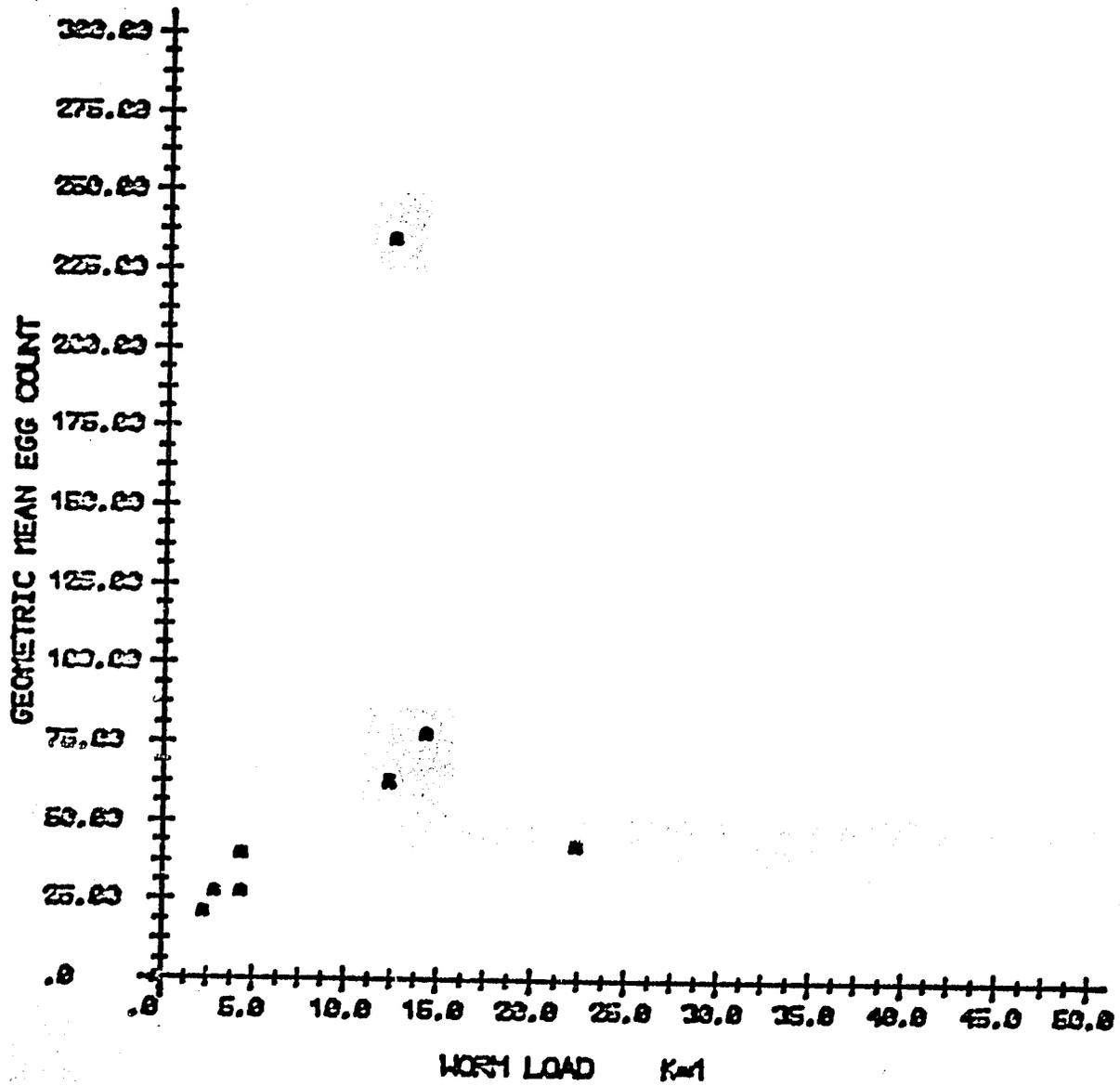
SECTOR 3 MALES AND FEMALES COMBINED



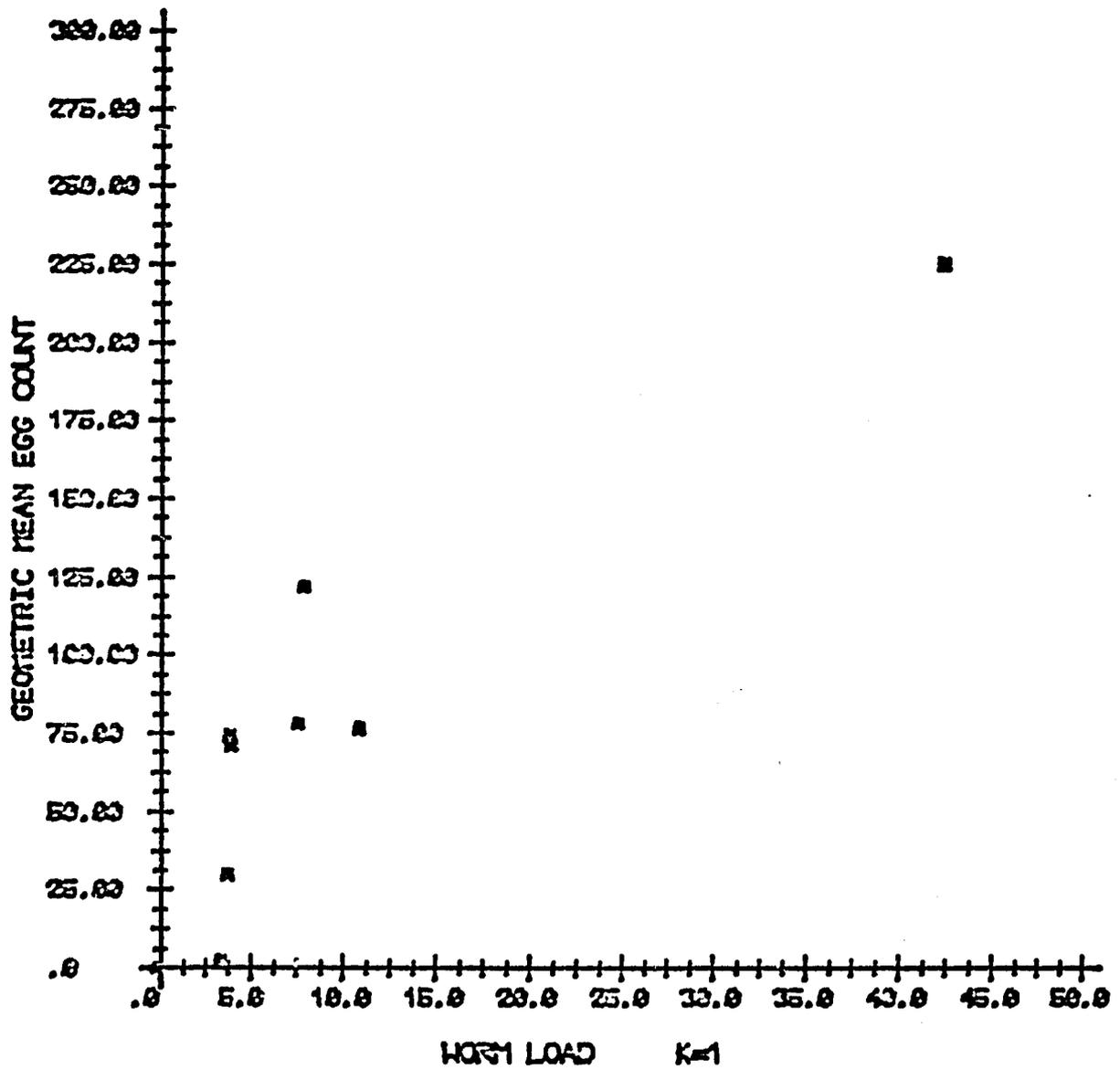
SECTOR 4      MALES AND FEMALES COMBINED



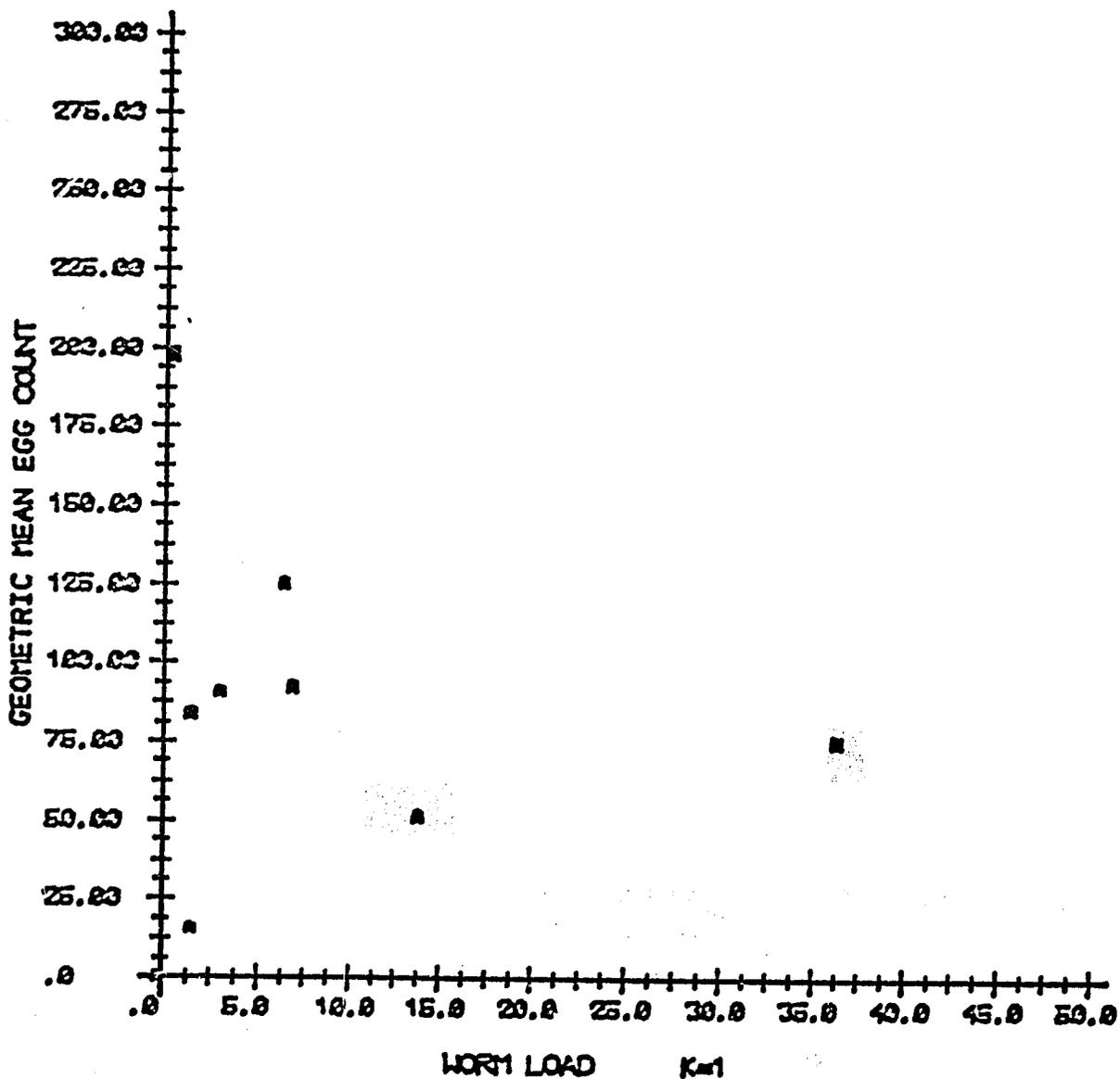
SECTOR 1 FEMALES



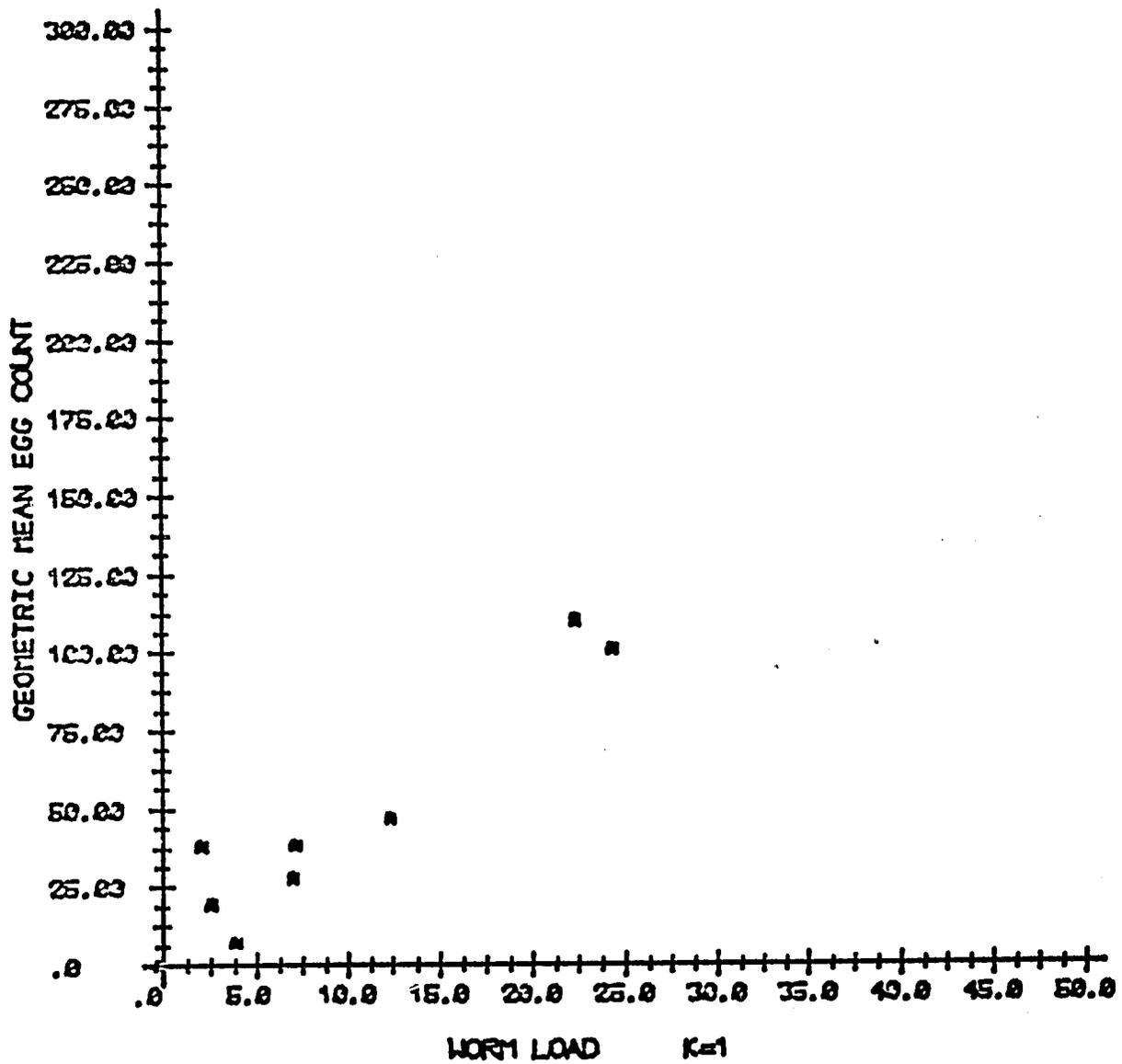
SECTOR 1 MALES



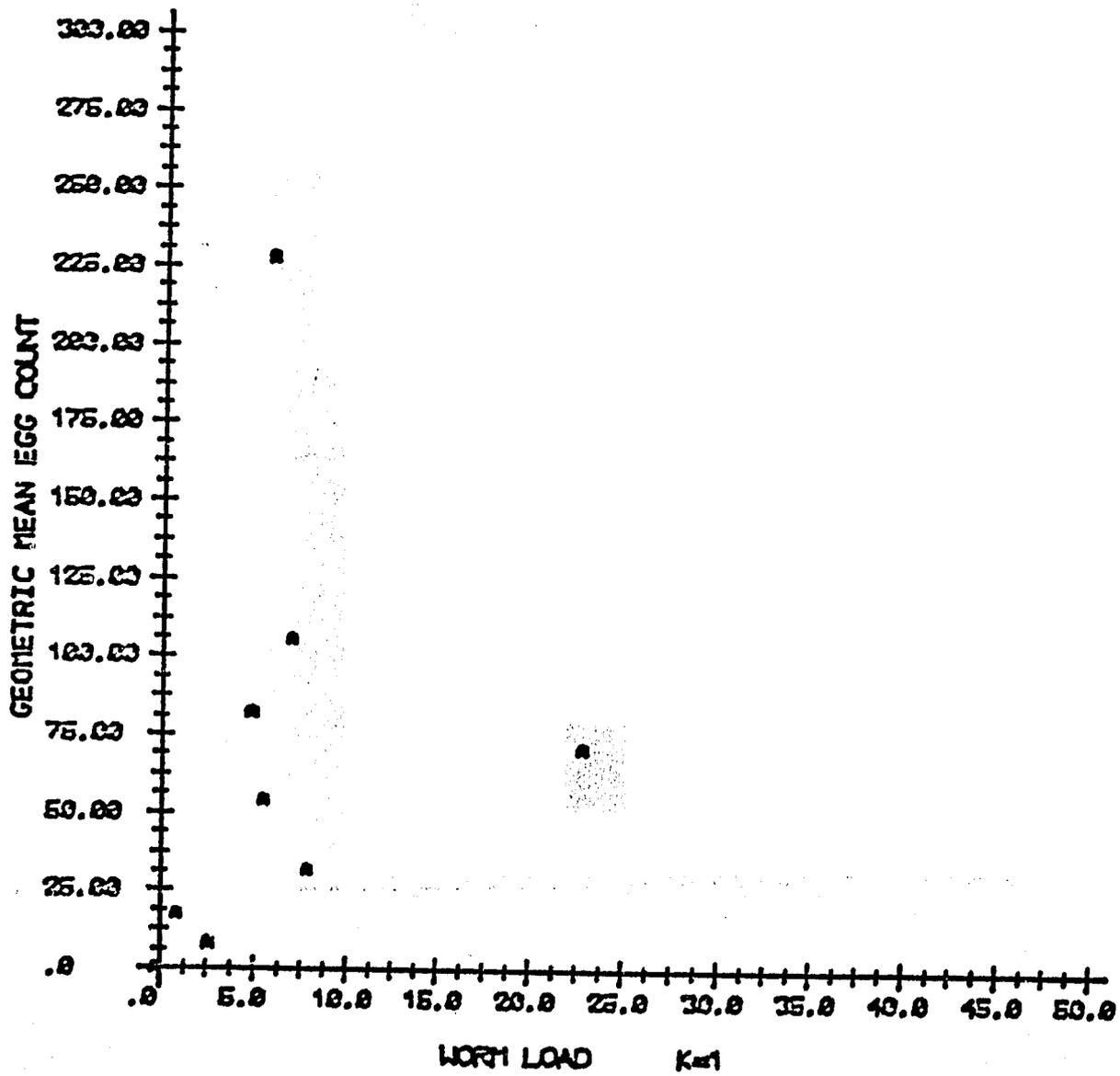
SECTOR 2 FEMALES



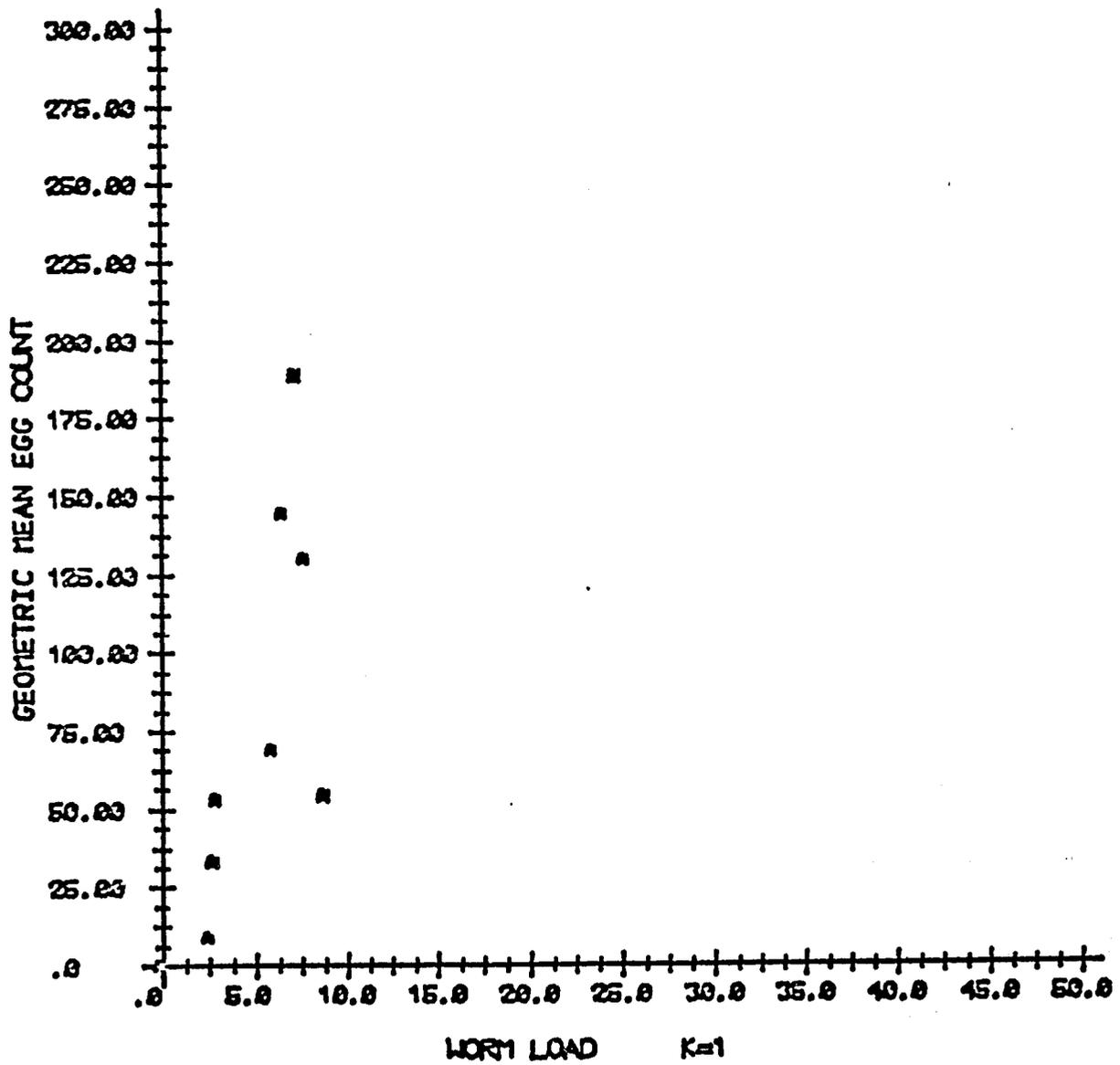
SECTOR 2      MALES



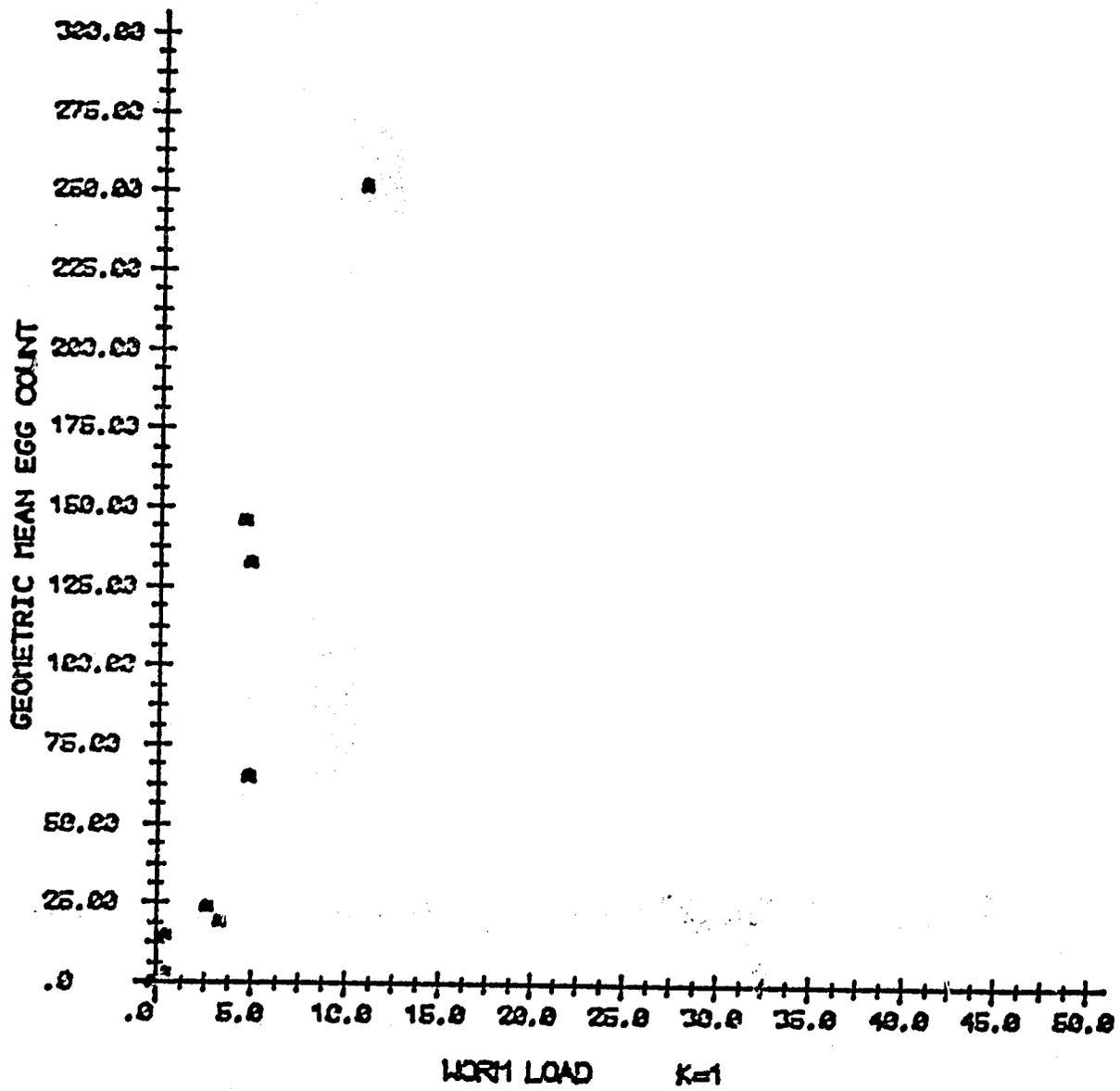
SECTOR 3 FEMALES



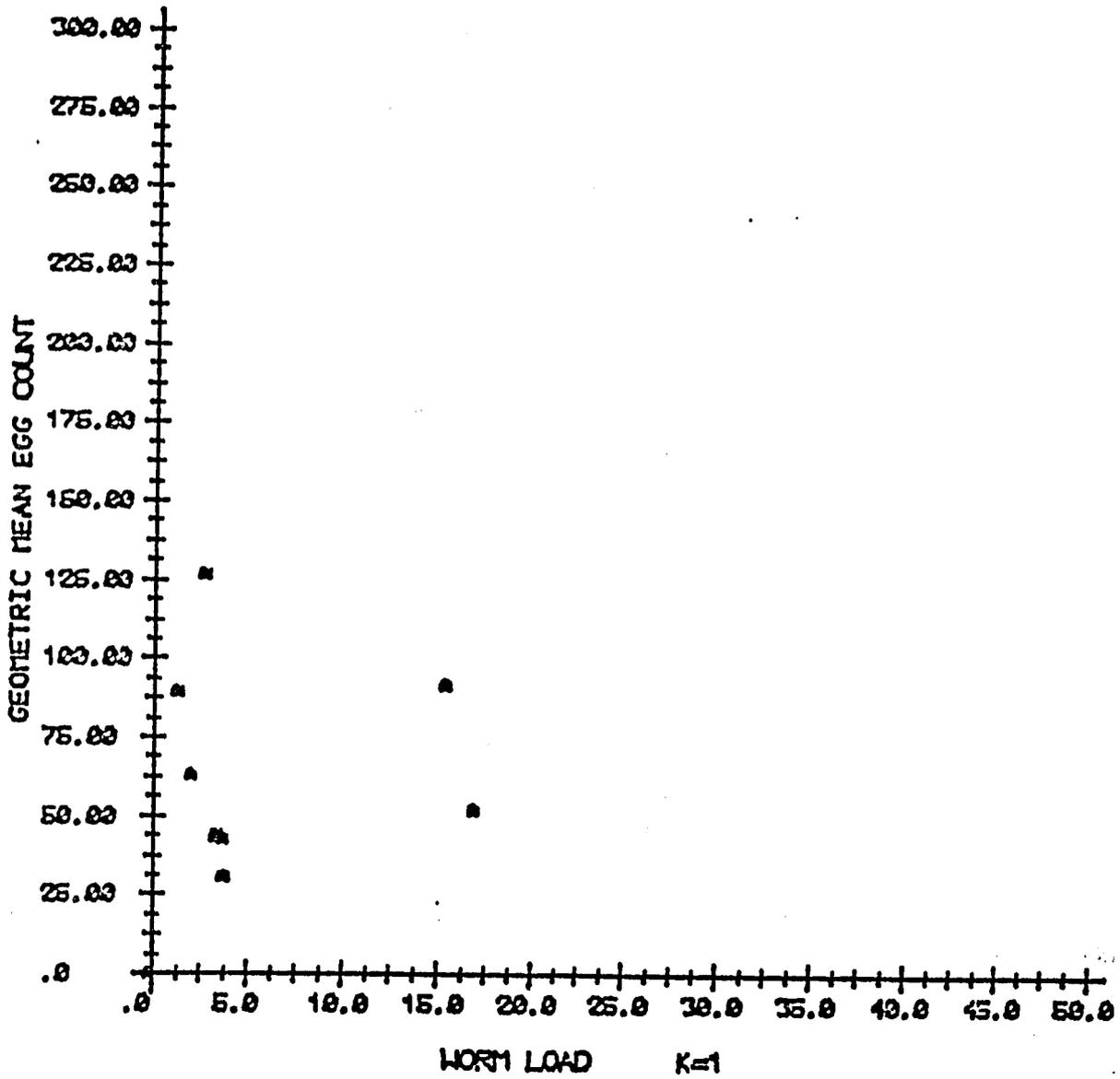
SECTOR 3 MALES



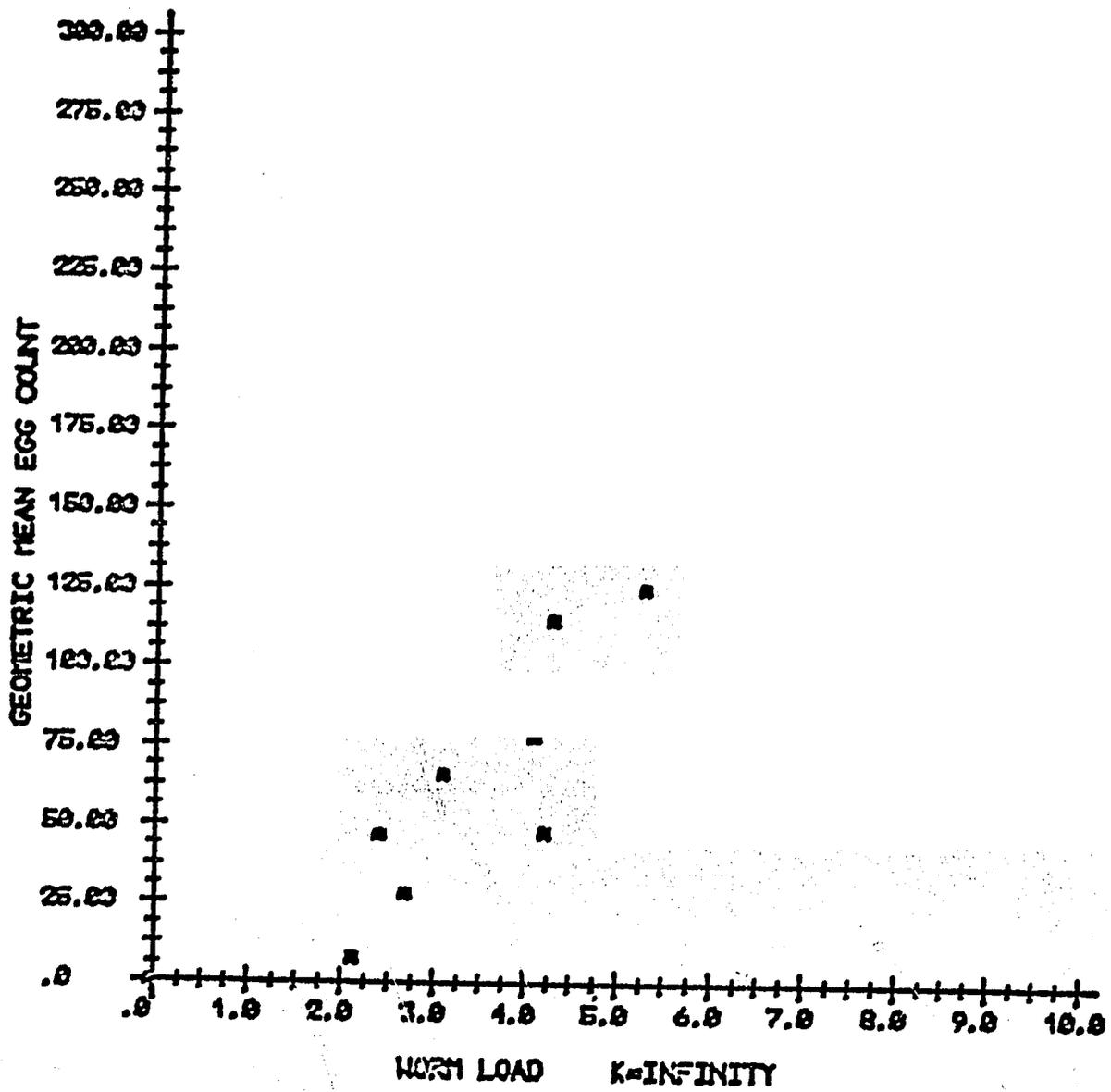
SECTOR 4 FEMALES



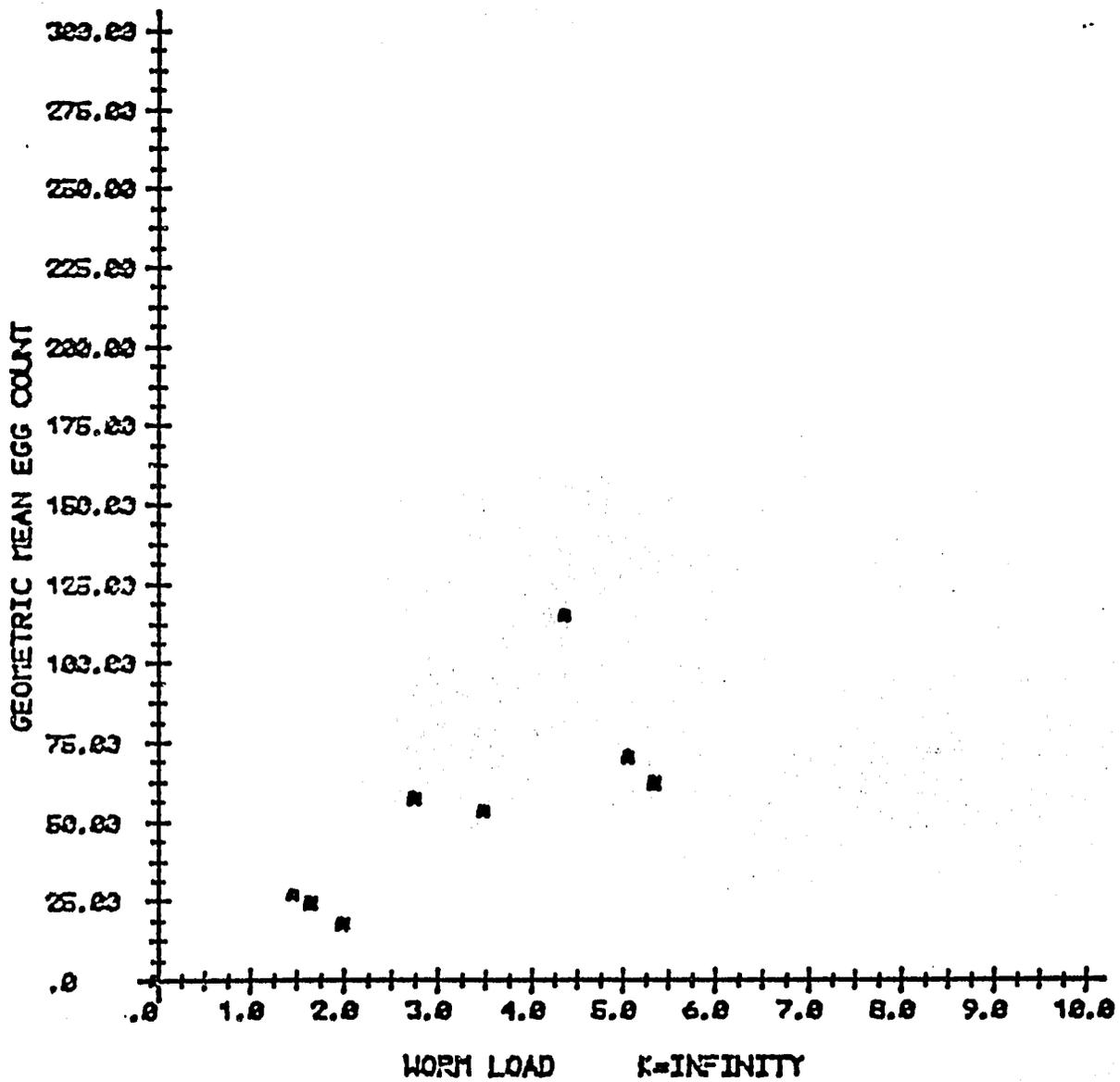
SECTOR 4 MALES



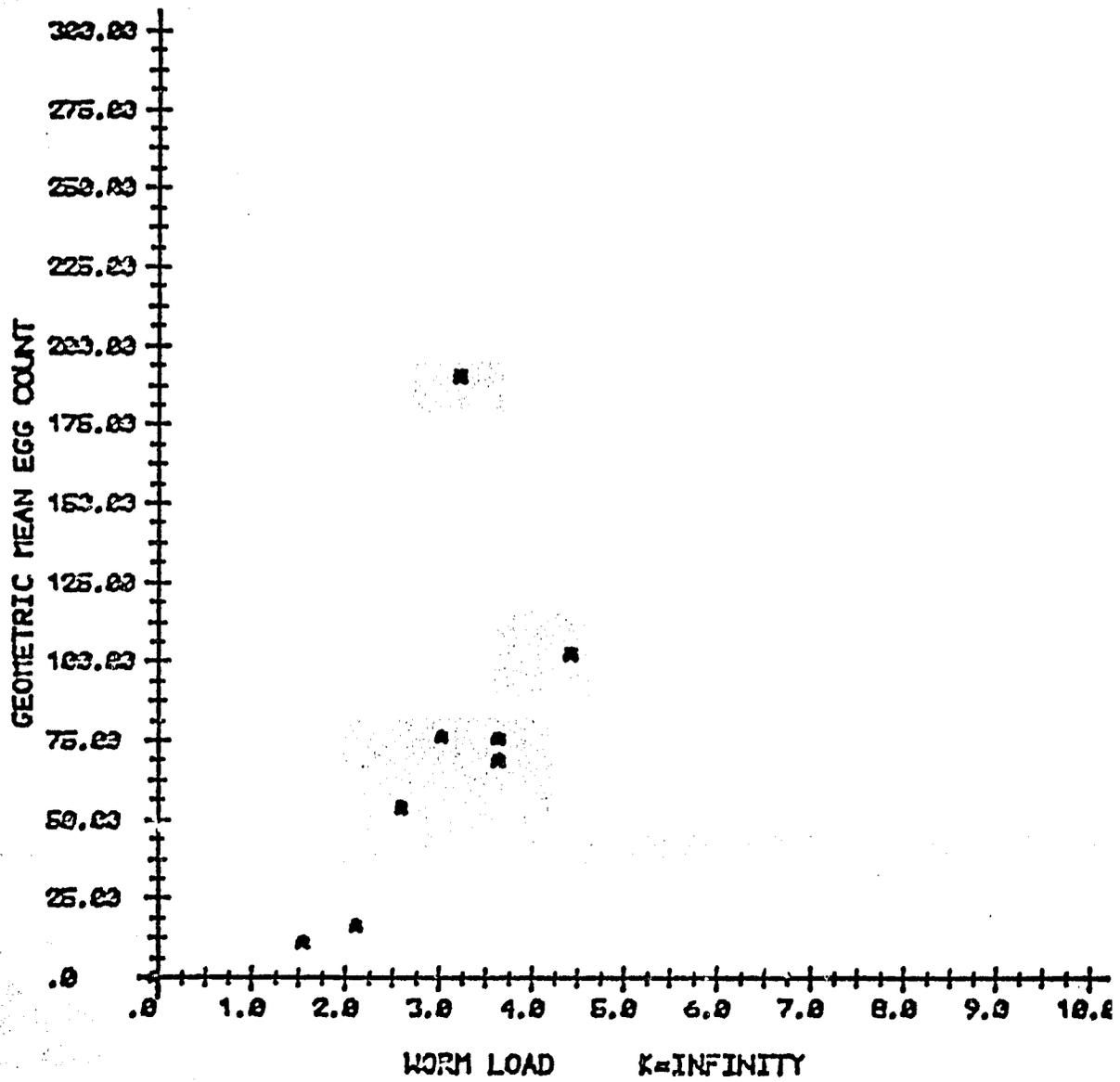
SECTOR 1 MALE AND FEMALE COMBINED



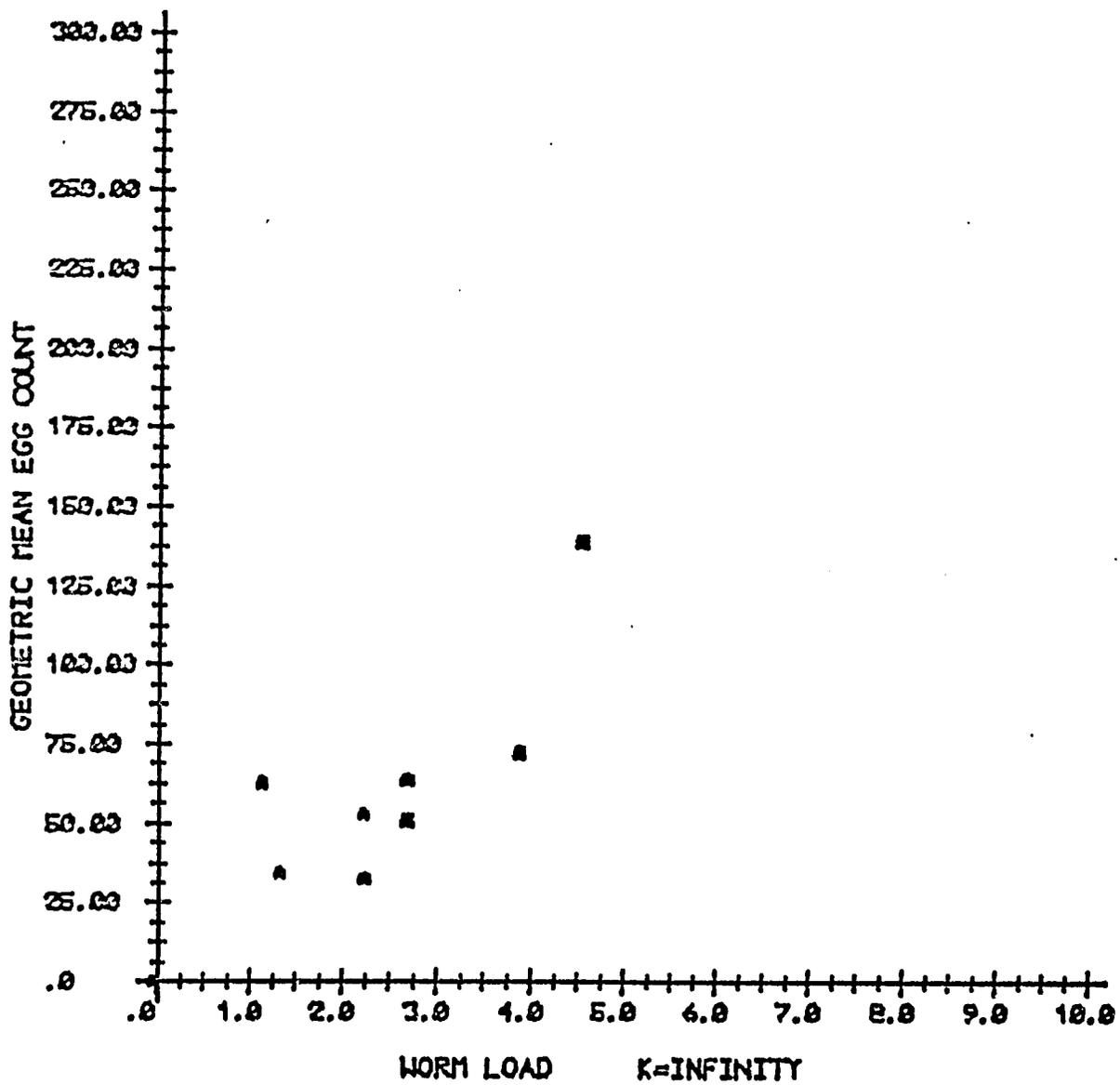
SECTOR 2 MALES AND FEMALES COMBINED



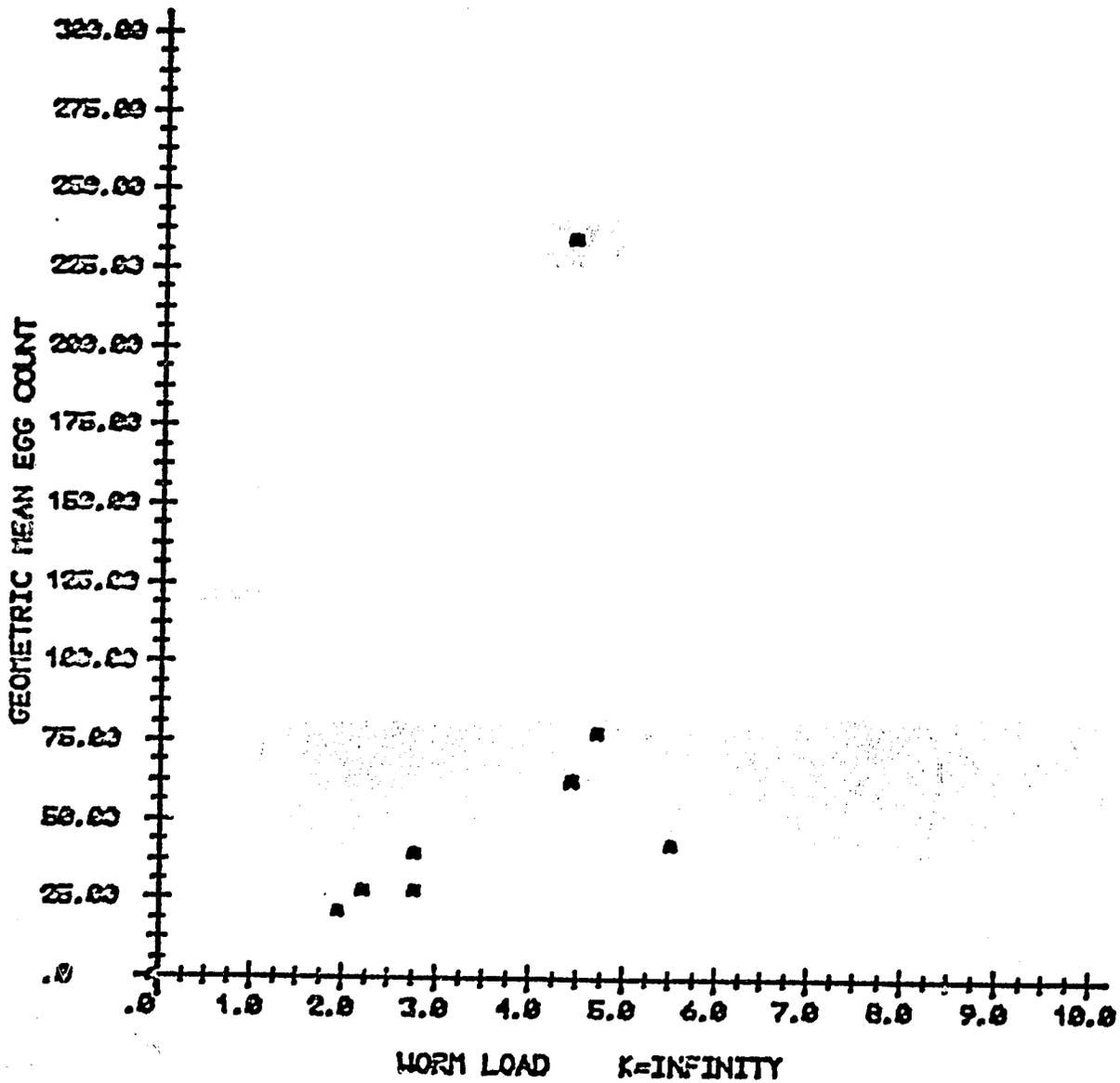
SECTOR 3 MALES AND FEMALES COMBINED



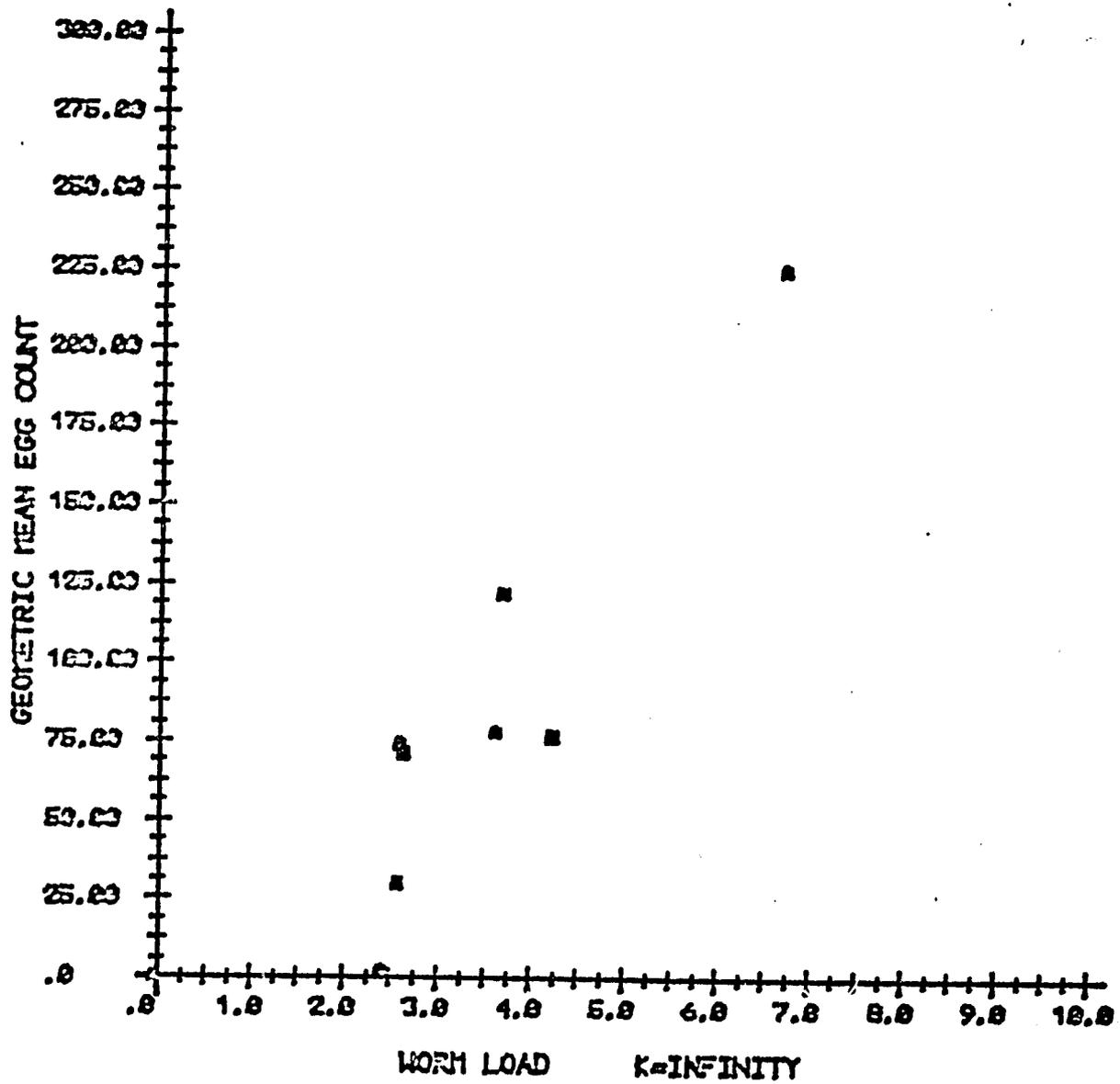
SECTOR 4 MALES AND FEMALES COMBINED



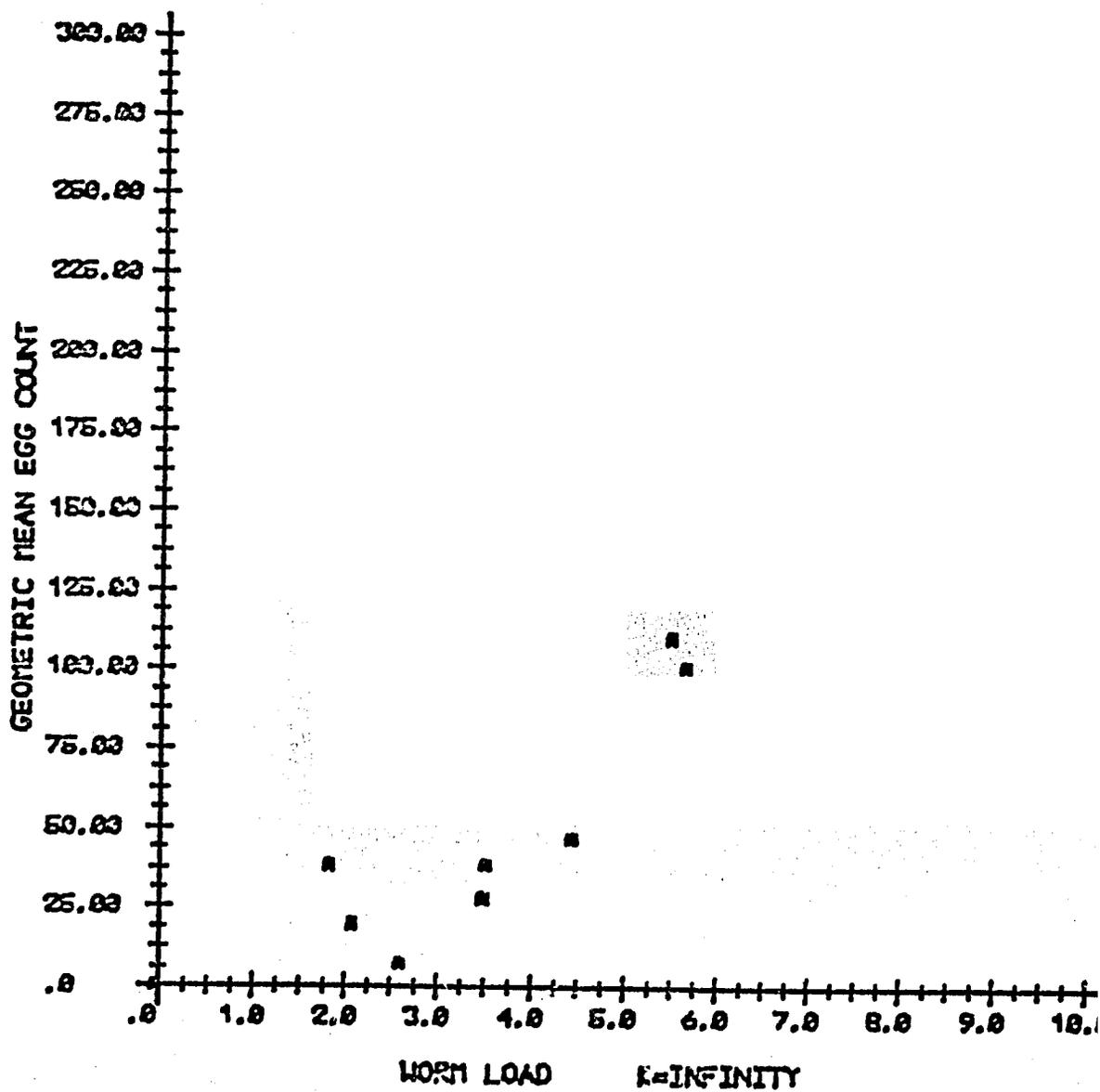
SECTOR 1 FEMALES



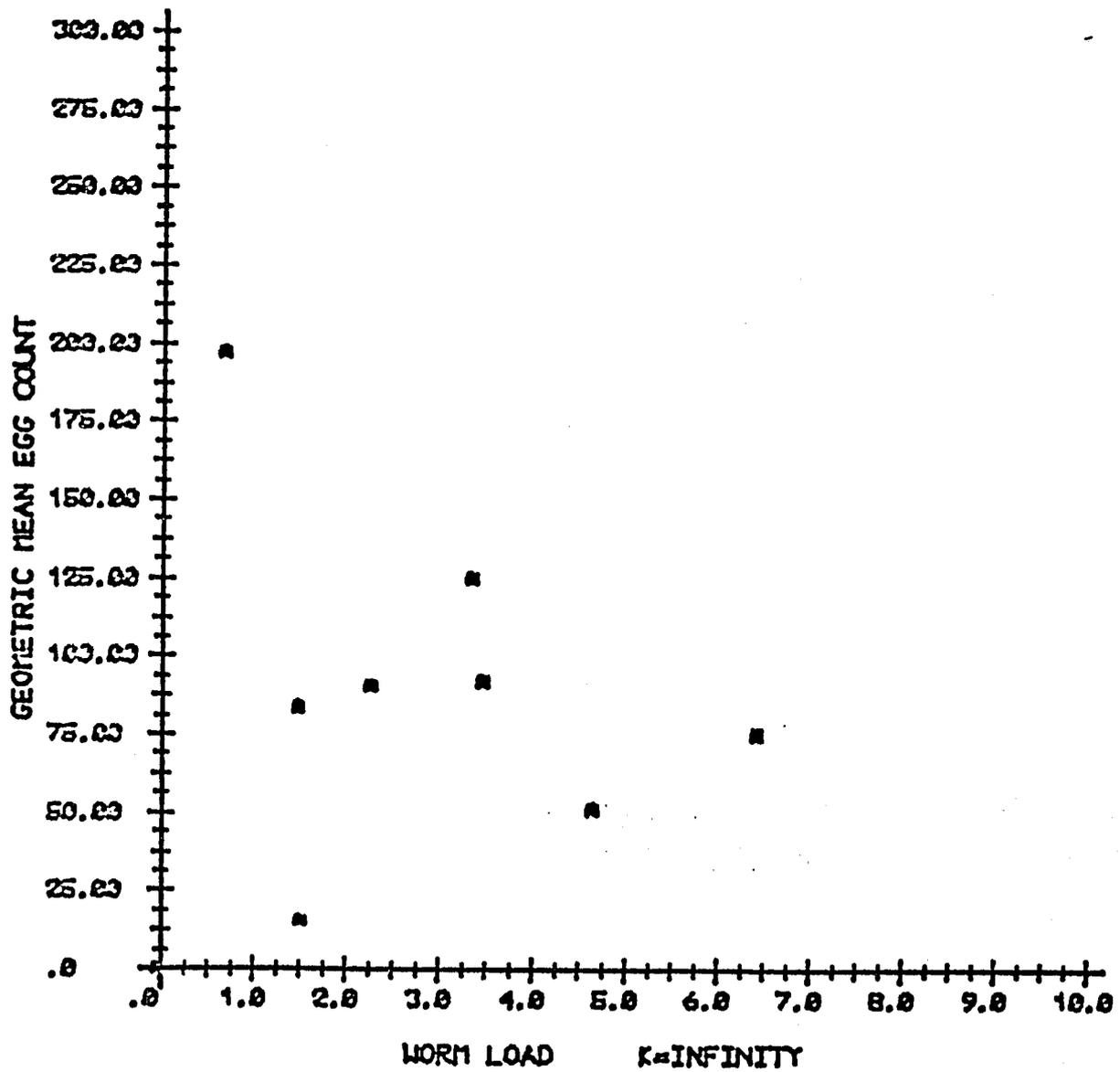
SECTOR 1 MALES



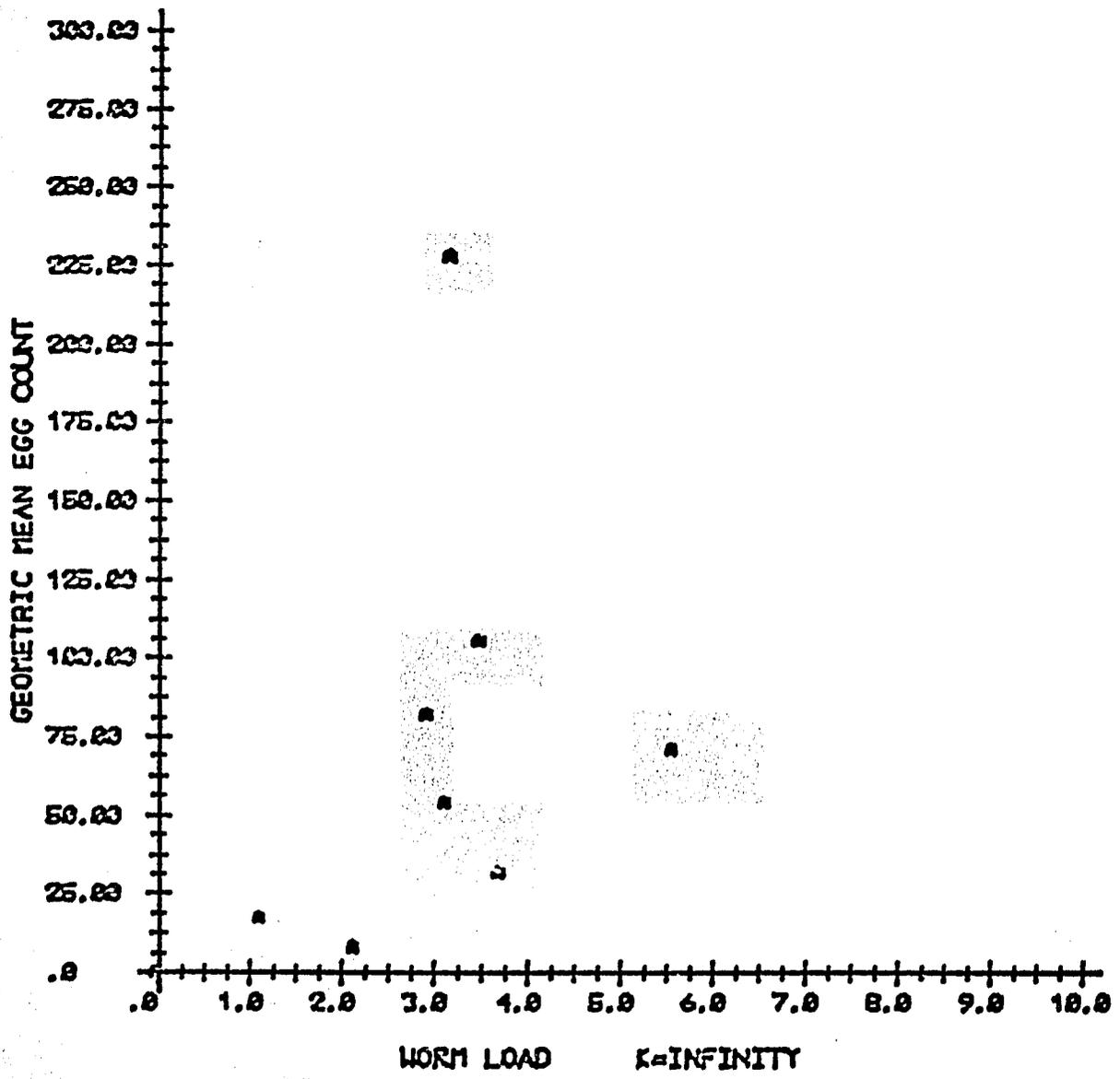
SECTOR 2 MALES



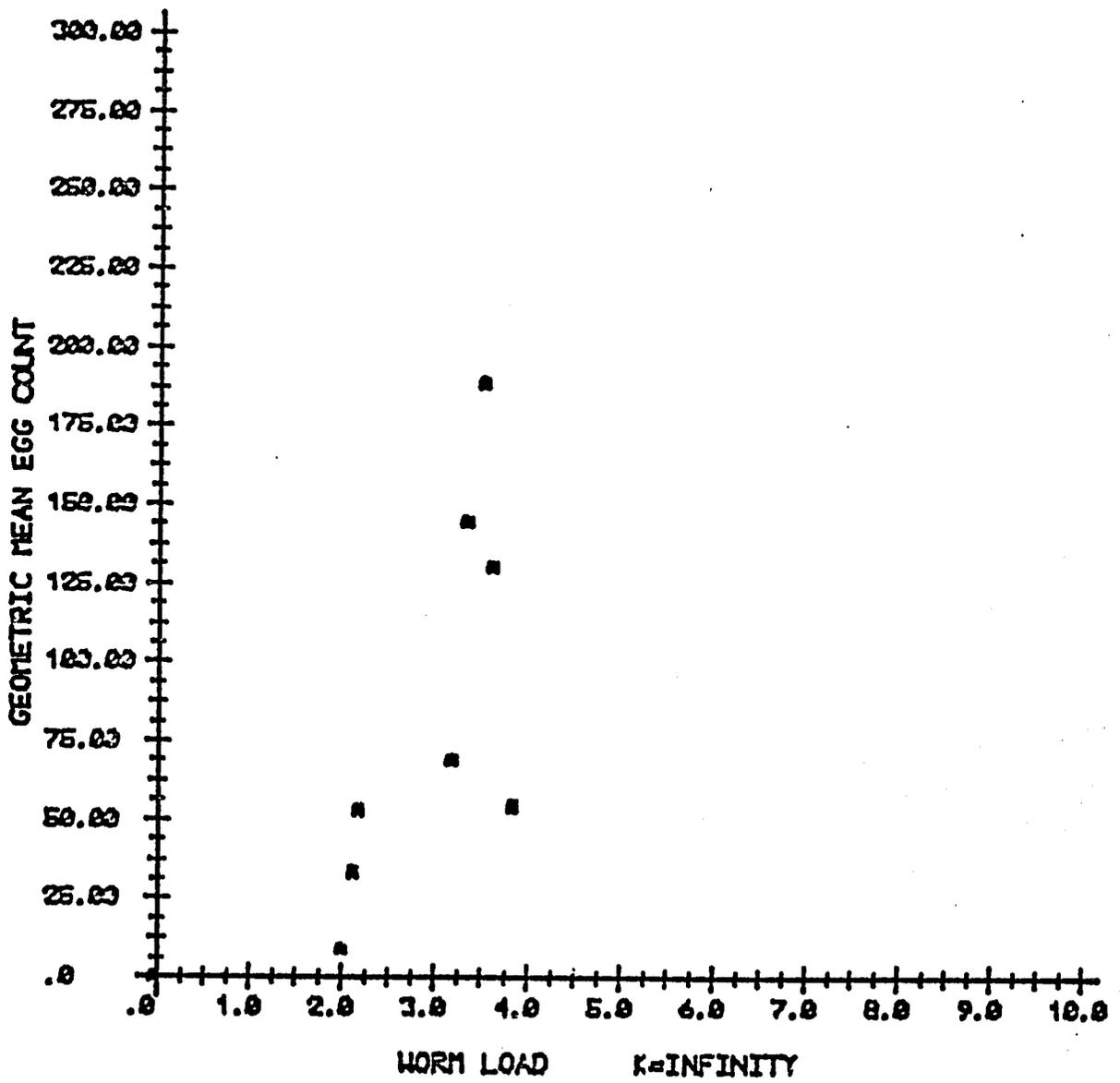
SECTOR 2 FEMALES



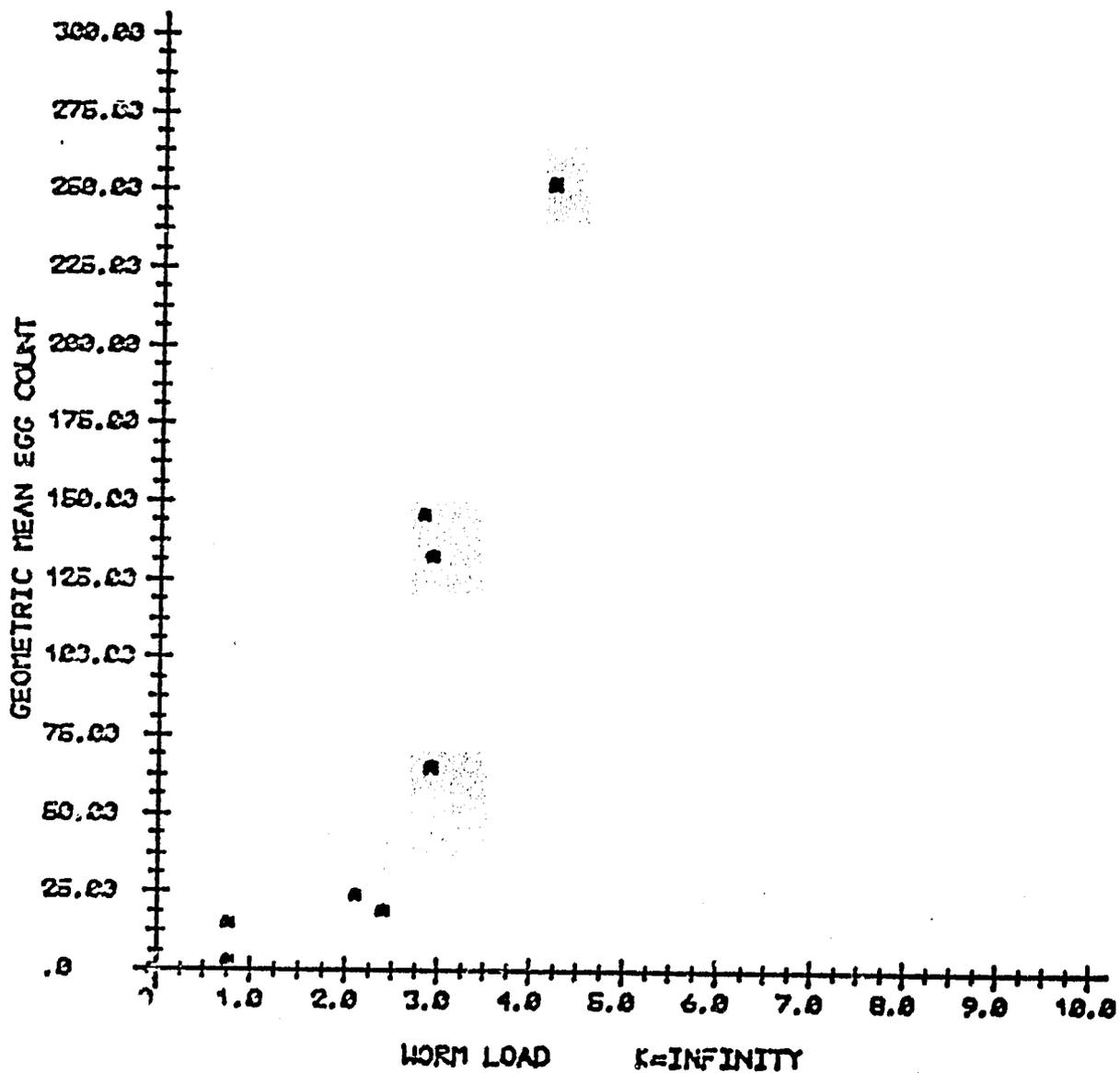
SECTOR 3 FEMALES



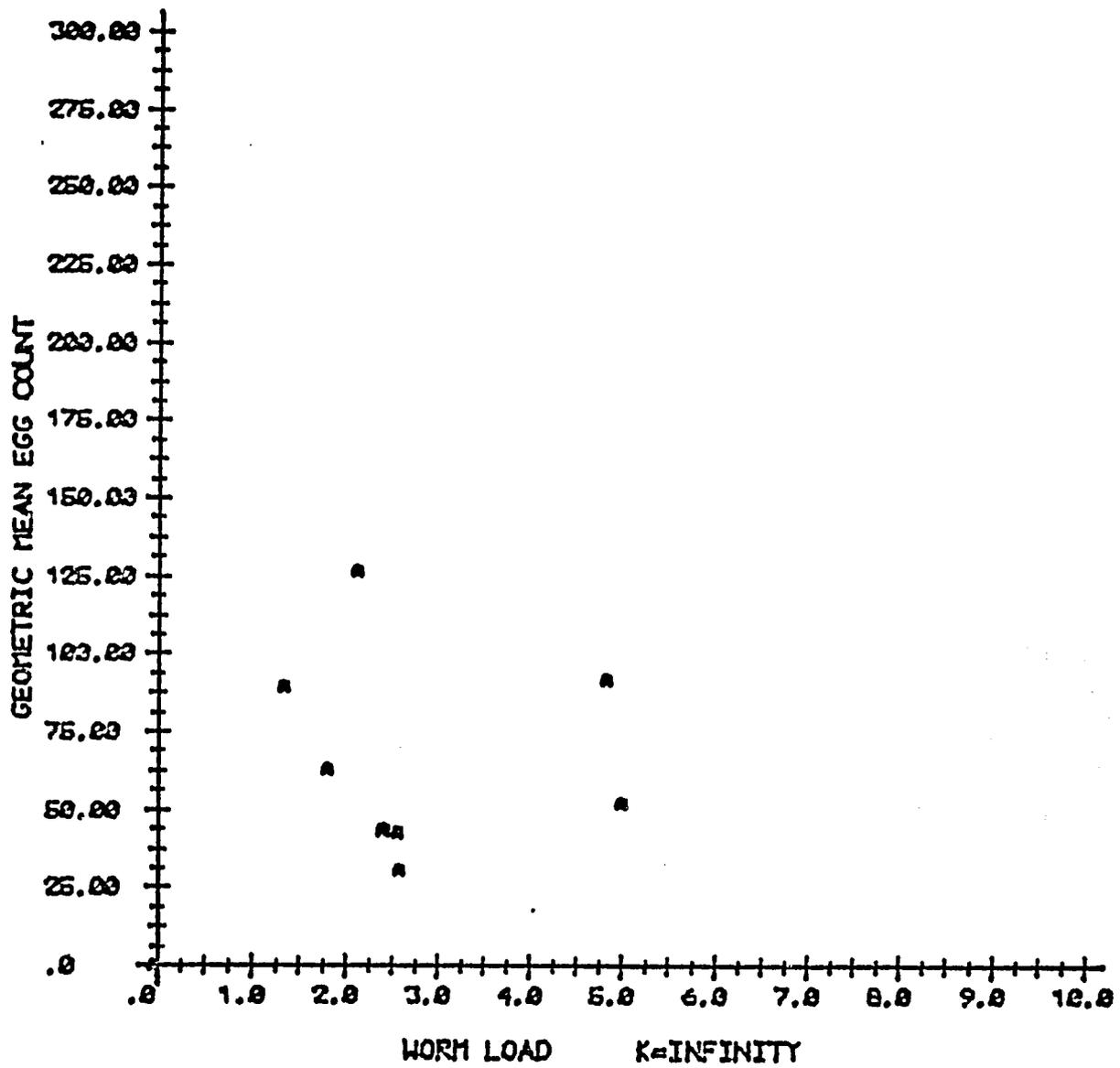
SECTOR 3 MALES



SECTOR 4 FEMALES



SECTOR 4 MALES



Appendix I

## Appendix I

### Processing of Tanzania 001 Data

In Appendix I we have listed the different stages in data processing undertaken for this study. In addition, we briefly refer to some questions associated with the data that relate to our use of them.

1. Stage 1 looked at raw data.

- a. Printed data in readable form. In order to use computer tape efficiently, the data were coded on the tape by the project in a very compact form. Before using the data in computations or listing it in tables, it was put into a less compact, more usable form via several computer programs.
- b. tried to interpret the meaning of each item of data.
- c. Resolved inconsistencies or vagueness of meanings of data. Coding sheets were a problem because of their incompleteness.
- d. Decided the relationship between separate files of data. Different files supposedly coded different data, but in one case, the 2 to 9-year-olds, the epidemiological data were in two separate files. One file contained data for all individuals in the project and the other file contained data for only 2 to 9-year-olds. However, there were discrepancies between supposedly identical data items for the same individual listed in the two different files. Moreover, some data were present in one file that were not present in the other file. The 2 to 9-year-old

file listed yearly egg counts for each individual, the entire population file listed such data at two-year intervals. A decision was later made to use the file containing the data for all individuals in the project; as a result, incidence for the 2 to 9-year-olds had to be computed over a two-year period.

- e. Made a list of unresolvable problems in interpreting raw data. Requested clarification from individuals involved in the original project.
  - f. Eventually, it became clear that there was no way to resolve the basic problem of the relationship of habitat to household without a trip to Geneva to view the project maps.
2. Stage 2 summarized raw data.
- a. Prevalence, egg count, migration distributions were calculated by sector, age, and sex for all individuals.
  - b. Incidence for 2 to 9-year-olds was calculated on the basis of the yearly data. The reexamination of individuals was not made at a fixed interval from the first one. Therefore, in calculating the incidence, a correction had to be made to standardize the interval.
  - c. Calculated geometric mean egg counts, and the worm loads predicted by various statistical distributions such as the negative binomial (with various coefficients) and the Poisson. These were plotted in various combinations to see whether the obtained relationships were similar to what would be predicted by theory. This was a preliminary step to using data other than the number of infected people in the regression equations which predict the incidence.

3. Stage 3 examined and duplicated maps in Geneva
4. State 4 read maps (time consuming). For each house we had to draw radii of various sizes. For each distance we read from the map which habitats were within the specified distance from each household. Looking at the computer output, we selected those habitats which were possible transmission sites. This was difficult to do because of the incomplete explanation on the coding sheets. These data were then keypunched on cards for input to the computer.
5. Stage 5 predicted incidence. Wrote data program to predict incidence from habitat volume, prevalence and egg count data. This program selects from the data on the tape the appropriate habitat data for each house, selects the individuals whose houses were plotted on the map (some data on the computer was for individuals living in houses not plotted on the maps), selects the 2 to 9-year-olds (since these are the only ones for which we can calculate incidence) and then matches individuals to the appropriate habitat data. From this we calculated for each age group in the 2 to 9-year category the total volume of habitat to which the group was exposed, the number infected, the mean egg counts and the geometric mean egg counts. This was used as input to the regression program which predicts incidence.
6. Stage 6 predicted prevalence. A computer program simulating the recursive prevalence equations was written. The input to this program was the age-specific incidence values calculated in Stage 5, the size of each age group and the age-specific volume or habitat.

For comparing prevalence predictions to observed data and also to summarize baseline data, we had to make some assumptions about the group in which to place individuals. There was some question in our minds from the data as to when "age" was recorded, at year of registration or at treatment. We assumed that age was recorded at treatment. Therefore, if someone had egg count data first reported for 1970, and was listed as 3 years old, we assumed that the individual was 3 years old in 1970. We feel comfortable with this assumption because in a number of instances we have 1968 egg count data but the individual was registered in 1969.

Appendix II

## Appendix II

Calculations and Assumptions for Including  
Different Terms in the Model

In Appendix II, we give the methods we used for calculating the terms in the model. In addition, we describe the decision criteria for inclusion of households and habitats. Data for individuals are used in the calculation of reversion and incidence rates only if egg counts are given for both 1968 and 1970 and data in computer file:

$$\text{Incidence} = \frac{\text{Number of people uninfected in 1968 but infected in 1970}}{\text{Number of people uninfected in 1968}}$$

$$\text{Reversion} = \frac{\text{Number of people infected in 1968 but uninfected in 1970}}{\text{Number of people infected in 1968}}$$

$$\text{Arithmetic mean of egg counts} = \frac{\sum_{i=1}^N \text{egg count of individual } i}{N}$$

$$\text{Geometric mean of egg counts} = \frac{\frac{1}{N} \sum_{i=1}^N \log (\text{egg count of individual } i)}{N}$$

Population increases were calculated assuming three-fourths of a percent increase per year. Therefore, for a two-year period, population (year 2) = 1.015 x population (year 0).

Migration (See discussion in text.)

Snail habitatsVolume (V)

When computing the volume of habitat, the habitat was treated as a sector of a sphere if length equaled width and as a sector of an ellipsoid otherwise. Letting L = length, W = width and D = depth, then:

$$\text{If } L = W, \text{ let } R = \sqrt{D^2 + L^2} / 4 \text{ Then } V = \pi R^2 D - \frac{\pi R^3}{3} + \frac{\pi (R-D)^3}{3}$$

$$\text{If } L = W, V = \frac{\pi L W D}{4}$$

To compute the volume of accessible habitat for each household, the volume of all the habitats at the given distance from the household were summed. To compute the age-specific habitat volume, each person in a particular age group was assigned the volume of his household, and the assigned volumes for all members of the age group were summed.

Perimeter (P)

When computing perimeter, the habitat was treated as a circle if length equaled width and as an ellipse otherwise:

$$\text{If } L = W, P = \pi L$$

$$\text{If } L \neq W, P = \pi \sqrt{\frac{L^2 + W^2}{2}}$$

The computation of the perimeter of accessible habitat for each household and each age group was similar to the corresponding volume measures.

## Decision Criteria with Reference to Households and Habitats

### Households

All households which could be located by number on the Sector IV map and for which data were available in the computer file were used. Out of the computer data file of 124 households, we used 117.

### Habitats

The decision criteria for habitats were more complicated. Not only did we need to use habitats located on the map by number and also in the computer data file, but we also had to decide which habitats were really transmission sites. The distance criterion described in the text was an assumption on our part about habitat accessibility. Next, we determined: 1) whether there were schistosomiasis carrying snails (in at least one of two surveys when the site was resurveyed); 2) whether the habitat was in fact dry (as indicated by the project under the item "is the water used"); and 3) the transmission potential of the habitat as determined by the project. Since there were many habitats that were originally surveyed by the project during 1967 and resurveyed in a different season in 1968, we could check a) to see if the habitat were snail-free consistently or if B. (P.) nasutus or other hosts ever appeared and b) to see if the habitat remained dry all year. The project had identified drinking ponds as unlikely transmission sites because users usually kept them

quite clean and clear of vegetation. Nonetheless, many drinking ponds were both used frequently and had B.(P.) nasutus so they were included. The decision criteria for determining which snail habitats to include may be summarized as follows:

1. Habitat must be numbered on map and have data in computer file.
2. Snails which transmit S. haematobium must be present or if two years of previous data, present at least in one year.
3. Transmission potential must be considered "probable" or "possible" by project. If considered "unlikely" or "equivocal" by project but appropriate snail species were present, the habitat was included.
4. Frequency of use must be "frequently," "very rarely," or "occasionally" with appropriate snail species in at least one survey.

The main criteria were presence or absence of data, presence or absence of S. haematobium bearing snails, transmission potential, frequency of use, and resurvey data.

Appendix III

## Appendix III

Suggested Statistical System for Schistosomiasis Control Projects  
(Preliminary Version)

The experiences of working with data from a number of schistosomiasis control projects has led us to consider additional data reporting methods to those already in use. The following appendix is included to offer some suggestions about the statistical aspects of control activities. The appendix is still in preliminary form and will be expanded in the final report.

Many reports have considered a variety of methods of organizing the collection of data and reporting results of schistosomiasis control projects (1). These reports have given careful thought to data recording forms, statistical methods for analyzing data, and techniques for collecting urine and fecal samples, counting eggs, estimating snail habitats, and organizing control activities (2).

Far less attention has been given to users of the information collected by a control project, that is, those persons not associated with the project who might be interested in using the collected data for cross-country or cross-project comparisons, testing models, developing evaluative methodologies, or simply deciding what the next control phase should be in the same area. It is essential to recognize that the data collection process is time consuming and takes much dedication on the part of the control project staff. They should certainly have

the opportunity to work with the data before external investigators. Nonetheless, these are not excuses for the wide variety of reporting mechanisms used over time by schistosomiasis control projects. Raw data and reports from many of the earlier projects are now lost to researchers, an unfortunate and possibly costly situation for the future of schistosomiasis control since the information from those projects might have assisted in the design of new projects.

A statistical system that might be quite useful for future schistosomiasis control projects has been developed and described by A. Mohapatra in a report to the World Bank on Family Planning Programs (3). In this he states: "An adequate service statistics system includes inter-linked subsystems of records, reports, data analysis, interpretation and feedback as its principal components" (4). The objectives of the program and resources available will define the data needed and the sophistication of such a system. Nonetheless, attention should be given to resource information (manpower, facilities, equipment, and financing), project activity statistics (staff activity reports, population information, follow-up reports), other health and vital statistics, and special surveys or studies.

The report should be organized so that input, activity, and output information is easily obtainable. The most efficient way to do this is to code the data collected and have them punched onto computer cards so that they are accessible and preserved permanently.

The frequency of reporting should be related to needs of the project staff for evaluating results, needs of the program supporter for review purposes, and capacity of data processing system.

The reports should contain uniform data reporting sheets so that report results may be easily compared. However, it is frequently necessary to modify data collected as the project continues. Thus, some flexibility is needed in the design of reporting forms, but the forms should maintain comparability over time.

#### Data Content and Presentation

Information usually collected by schistosomiasis control projects includes: number of persons examined, number infected, egg counts (occasionally), incidence values for children, snail species, types of snail habitats, snail population characteristics (seasonal variation, susceptibility to different molluscicides, etc.), number of persons treated, post-treatment egg counts, drug dosages, molluscicide formulation, and mollusciciding success (success defined as no observed snails in treated habitats). A scattering of other data may include: maps of habitats and households, seasonal climatic patterns, topography, reversion rates, population movements, water contact patterns (ranging widely in scope), cercarial studies, animal reservoir infection, clinical studies (range of severity of infection), overall census information for the initial year of the project, and control cost data. Very rarely is information given on living conditions, employment conditions, general economic status of region, vital statistics such as birth and death rates, and prevalence of other diseases. The most discouraging aspect is that usually one set of data is collected without a complementary set, for example incidence rates without water contact patterns. Most project staffs, despite their best intentions, collect data on only a subset of the items. Manpower and financial resources

are often serious constraints to the collection of a complete set of data. It should be mentioned that another reason is the lack of consideration as to who else might use the data in the future. Thus, many data are useless because the counterpart information is lacking.

In Tables III-1 and III-2 are listed information that modelers might want. These two tables together represent extreme needs. Indeed, the information in Table III-1 would be very difficult to collect in the field. If one relies on laboratory results for filling in this information, one then has to make assumptions about the transferability of laboratory results to field situations. The information in Table III-2 represents the other extreme: field data which may be too costly to collect. Some set of data in between these extremes is more appropriate for use in predictive models. Because of resource constraints, priorities need to be established about what information is critical for assisting in decision-making. We present in Table III-3 a preliminary listing of priority data required for predictive modeling and evaluation purposes. We will refine this table for the final report.

The units used are critical. Indeed, for this reason, it is suggested that data be reported on an individual basis (at least semi-yearly). Even if groupings are to be made, such as in calculating incidence or reversion rates, each age should be reported separately so that the data user may make his own judgment about groupings.

Related data should be presented in comparable units. If habitats are measured in feet, the molluscicide formulation should be calculated on a per foot or cubic foot basis. If maps contain individual households

Table III-1. Rate Information Needed for Transmission Models, from Hairston\*

---

- I. Survival of Eggs that Reach the Excreta of the Definitive Host
  - A. Hatchability
  - B. Survival of hatchable eggs
    - 1. Probability of being deposited in water containing snail hosts
    - 2. Survival of miracidia while free in the water
    - 3. Probability of locating and penetrating a host snail
    - 4. Probability of establishing an infection after penetration
- II. Survival of, and Reproduction by, Larval Stages in the Snail Host
  - A. Survival of mother sporocysts
  - B. Reproduction by mother sporocysts
  - C. Survival of daughter sporocysts
  - D. Reproduction by daughter sporocysts
  - E. Survival of snail host after the start of production of cercariae
  - F. Probability of cercariae escaping from the snail host
- III. Survival of Cercariae
  - A. Survival of cercariae while free in the water
  - B. Probability of coming into contact with a definitive host
  - C. Probability of successful penetration of the skin
- IV. Survival of Young Schistosomes in the Definitive Host
  - A. Survival during migration through skin, lymph glands and lungs
  - B. Probability of finding the appropriate blood vessels
  - C. Survival in appropriate blood vessels until maturity

Table III-1. Rate Information Needed for Transmission Models, from Hairston\* (continued)

---

- V. Survival of, and Reproduction by, Adult Worms
    - A. Probability of locating a mate
    - B. Reproduction by the female schistosome
    - C. Survival of adult worms
    - D. Survival of the definitive host during infection
  - VI. Probability that Eggs will Reach the Excreta of the Definitive Host
- 

\*Hairston, N., 1973. "The Dynamics of Transmission" in N. Ansari, ed. Epidemiology and Control of Schistosomiasis (Baltimore, University Park Press), pp. 279.

Table III-2. Data Necessary for Representative Transmission Models, from Rosenfield\* \*

Data Requirements	Population	Environment
Humans	Age/sex distribution Birth rate/age/sex Death rate/age/sex Prevalence/age/sex Incidence/age/sex Migration rate/age/sex	Proximity of homes to water Sanitary conditions Frequency of contact with infected water/age/sex Nutritional and health level/age/sex Health facilities
Snails	Age distribution Birth rate/age/sex Death rate/age/sex Species Density of infected snails	Dimensions of available habitats Food supply Velocity of water Light conditions Pollution of water (degree and kinds) Chemical conditions of water Seasonal variations of population
Schistosomes	Probabilities of pairing, egg hatching, miracidia and cercariae penetration Mean worm load in humans/related to number of eggs passed Reproductive rate in humans Reproductive rate in snails Existence of compensatory mortality* Species in humans and reservoir hosts	Effects of chemotherapy/immunity in humans Effects of mollusciciding/immunity in snails Existence of break-point below which the probability of parasites mating is so low that the disease cycle is broken

\*Reaction of parasite population to environmental changes, if control measures eliminate large numbers of parasites through eradication of snails compensatory mechanisms may result in higher reproduction rates of the parasite in man. Thus, the disease may be as prevalent as before controls.

\*\*Rosenfield, P. L. 1975. Schistosomiasis. In: Parasitology, 2nd ed., pp. 1-10.

Table III-3. Preliminary Priority Listing of Data Needed for Predictive Modeling and Project Evaluation: A Modeler's Viewpoint

---

Epidemiological

Total numbers in each age group  
 Number examined/age/sex on an individual basis  
 Number infected/age/sex on an individual basis  
 Number eggs passed/unit urine or feces/age/sex on an individual basis  
 Map of household location in relation to types of snail habitats  
 Snail species  
 Population migration data  
 Transmission sites (which habitats play a role in transmission)  
 Human contact with transmission sites/age/sex with specific transmission sites identified for each individual  
  
 Incidence and reversion rates/age/sex  
 Characteristics of snail habitats (e.g., dimensions, vegetation, seasonal variation)

Environmental

Rainfall patterns  
 Topographic maps

Controls

Date of chemotherapy, reduction in egg counts/age/sex on an individual basis  
 Date of mollusciciding of individual habitat basis  
 Success/failure of treatment for individual habitats

Economic

Employment activities/age/sex  
 Domestic activities/age/sex  
 Wage rates for control activities (professional and other labor)  
 Equipment and material costs for controls

---

or habitats, these sites should be easily identified, for example, by number so they can be related to information in tables or data file. Maps should be labeled; that is, each map should have a complete key with scale on it. The scale on project maps should be comparable to that on topographic maps. Each project map should be made to the same scale when possible or else scaled down or up proportionately (not free-hand) so that maps of different scales may be compared. It is imperative that the same information be collected on a pre- and post-treatment basis so that control measure effectiveness may be assessed.

## Appendix III

## References

1. WHO Expert Committee on Bilharziasis, 1965. Third Report, WHO Technical Report Series No. 299, WHO, Geneva; N. Ansari, ed. 1973, Epidemiology and Control of Schistosomiasis (Baltimore, University Park Press), 752 pp.
2. Oliveir, L. J. and K. Uemura, 1973, Technologies, Statistical Methods and Recording Forms in N. Ansari, ed. Epidemiology and Control of Schistosomiasis (Baltimore, University Park Press), pp. 620-748.
3. Mohapatra, P. S., 1977. Measuring the Performance of Family Planning Programs. World Bank Staff Working Paper No. 257, World Bank, Washington, 59 pp.
4. Ibid, p. 40.

IV -1

Appendix IV

## Appendix IV

This appendix contains a summary of the main consultations we have had for this project.

1. Meeting with Dr. Fergus McCullough, Former Director, Tanzania 001, now with World Health Organization, Geneva. Dr. Rosenfield and Gestrin met with Dr. McCullough from May 2 through May 4, 1978 in Geneva. The purpose of the trip was twofold:
  - a) Dr. McCullough had the Tanzania 001 maps which contained information necessary for the modeling attempts. Since there were no copies of the maps, he was reluctant to ship them to the U. S. but would let us copy them in Geneva. We had complete access to the maps there and were able to copy all the maps. In addition, there were many questions regarding interpretations of the maps which only Dr. McCullough could clarify.
  - b) Our second purpose was to obtain answers to data coding questions and others that had arisen as we worked with the data. Dr. McCullough and Mr. Hubert Dixon (who had put the data on tape in Geneva in 1976) met with us for several hours each day to go through these questions and other assumptions we had made. Some questions, especially regarding coding, could not be answered. We have thus contacted Dr. Syvend Christiansen on the advise of Dr. McCullough. In addition, we talked about cost data with

Dr. McCullough and he stressed the fact that the project was a practical, field-oriented one; thus, he thought a strict cost-effectiveness analysis might not be applicable. We are pursuing that concern in the next stage of the project.

The consultation was very important both to obtain data and answers to questions and, perhaps as importantly, to assure Dr. McCullough of our desire to keep him informed of our progress at every stage of the project. (We also briefly met with Dr. Rikk Davis and Dr. T. Lipes to inform them of our work on the project.)

2. The second consultation was with Dr. David Bradley, Director, Ross Institute of Tropical Hygiene, London, in person on May 4, by several telephone conversations, and by correspondence. Dr. Bradley, one of the most prominent authorities in schistosomiasis epidemiology, has been of great assistance in advising us on several phases of modeling in general and specifically with regard to Tanzania 001. In London, we discussed the egg count distribution questions in depth, especially the rationale for using the negative binomial distribution. We also discussed the interpretation of egg count data in relationship to acquired immunity. We also discussed the map data from Tanzania 001 and the possibility of determining habitat accessibility and water contact patterns for different age groups. Dr. Bradley very kindly offered to try to obtain aerial photographs of the project area from Mwanza, and also expressed a willingness to review our reports and results.

3. We have also consulted informally: Dr. Robert May, Princeton University, about egg count distribution; Drs. Leonard Berry and Richard Ford, Clark University, about Tanzania-living habits, and water supplies in Misungwi (Dr. Ford purchased some topographic maps of the area for us when he was in Tanzania), and general status of health in Tanzania; Dr. Fred Golladay, World Bank, about general status of health and water supplies in Tanzania; Dr. Richard Smith, United States Geological Survey, about modeling techniques for handling of egg count and migration data; Dr. Kerry Smith, Resources for the Future, for statistical comparison of Poisson distribution and negative binomial, and for migration data modeling suggestions; and Dr. Clifford Russell, Resources for the Future, for migration modeling, especially use of Markov chain process.

## ACKNOWLEDGMENTS

The authors acknowledge with gratitude the support of the Agency for International Development (AID) for this research. Several persons from AID deserve special thanks for their efforts and advice: Mr. James Thomson, Dr. Lee Howard, Mr. Dale Swisher (now retired), Dr. Joe Stockard, Dr. Miloslav Rechcigl, Dr. James Erickson, and Dr. Floyd O'Quinn.

The cooperation of the World Health Organization was essential for the completion of the study and is greatly appreciated. The authors are especially grateful to Dr. Andrew Davis, Dr. Fergus McCullough, Dr. V. Eyakuze, Mr. Herbert Dixon, and Dr. Svend Christensen for their assistance.

The authors would like to thank the following persons for their thoughtful suggestions regarding transmission modeling: Dr. David Bradley, Director, Ross Institute of Tropical Hygiene; Dr. Richard Smith, United States Geological Survey; and Dr. Robert May and Mrs. Joan Aron, Department of Biology, Princeton University. Dr. Richard Ford, Clark University provided generous advice and information on Tanzania.

Colleagues at Resources for the Future generously provided their time, effort, and knowledgeable advice, especially Dr. Emery Castle, Dr. Ken Frederick, Dr. Clifford Russell, Mr. Blair Bower, Dr. Kerry Smith, and Mrs. Louanne Sawyer. The support and assistance of the entire Computer Center staff of the Brookings Institution are also greatly appreciated.

The authors are especially grateful to Mrs. Loretta Burgess, Mrs. Sylvia Steadham, Mr. John Mankin, Mrs. Peggy White, and Ms. Lorraine Van Dine for their patient and careful typing of many drafts.

Responsibility for any errors lies, of course, with the authors.