Stochastic modelling of the transmission of endemic malaria with analysis of field data

Bakheit, Charles Saki

STOCHASTIC MODELLING OF THE TRANSMISSION OF ENDEMIC MALARIA WITH ANALYSIS OF FIELD DATA

by

Charles Saki Bakheit. M. Math. (Waterloo)

Thesis submitted for the degree of
Doctor of Philosophy
in the
University of Durham

Department of Mathematics
University of Durham
England
1982

PREFACE

The work presented in this thesis was carried out in the Department of Mathematics at the University of Durham, between September 1978 and July 1982, under the joint supervision of Drs. Allan Seheult and Peter J. Green.

The material contained in this thesis has not been submitted for any degree in this or any other university. No originality is claimed for the first three chapters except where indicated. Most of the applications in chapters 4, 5 and 6 are believed to be new. Materials used from other sources are appropriately referenced. Some of the algorithms used, especially in section 4 of Chapter 5, were kindly provided by my co-supervisor, Dr. Peter J. Green.

May I express my sincere gratitude to Drs. Allan Seheult and Peter J. Green for their guidance, continued encouragement and many helpful suggestions throughout the course of this work. My thanks also go to Dr. Eluzai Hakim-Abe, of the University of Juba (Sudan) Students Clinic, for reading through Chapter 1 and offering useful suggestions.

The World Health Organization Division of Malaria and other Parasitic Diseases provided me with tapes of the raw data from their Garki project. I was also later sent a free copy of the monograph by Molineax and Gramiccia (1980) after its publication. These services are gratefully acknowledged.

I also wish to thank the British Council for their financial support, and their representative, Miss E.M.Miller, for facilitating my contacts with the Council.

Many thanks to Mrs. Margaret Bell for her patience and superb typing.

Finally, to my wife, Saida, who had to forgo her honeymoon and has spent many lonely hours in the last three years, my deep appreciation.

ABSTRACT

A critical review of some current malaria models is given in which a new model of superinfection is presented. An alternative malaria model, partly stochastic and party deterministic, is then proposed and results of the simulation of the model are discussed. Simplified versions of the model are used to analyse longitudinal survey data from a World Health Organization malaria project carried out in Northern Nigeria.

INTRODUCTION

The last 80 years have seen much work devoted to the quantitative aspects of the dynamics of malaria. This work has produced models that are largely deterministic, and are developments and extensions of the pioneering model of Ross (1911). One major criticism of these models is that they emphasize the host-vector relationship almost to the total neglect of the host-parasite interaction (Najera, 1974).

In this thesis, our main objective is to propose a partially stochastic framework of malaria infection aimed at including some salient features of the interactions between the host and parasite populations. We attempt to achieve the latter by incorporating assumptions based on known empirical results about the host-parasite relationship.

In order to gain understanding of the disease and appreciate the complexity of its biology, we give some epidemiological background in Chapter 1.

In Chapter 2, we give a critical review of some of the better known deterministic models. We also propose a new model of the phenomenon of superinfection.

Chapter 3 contains a summary of some results of semi-Markov processes to be used in subsequent chapters. One of these results is believed to be new.

Chapters 4, 5, and 6 constitute the main body of the thesis. In Chapter 4 we formulate a hybrid semi-Markov model to describe the course of the disease in an individual from when he is infected to when he is susceptible again. Attempts

are made to formulate the distributions of times in each stage of the disease so that they reflect some specific biological features and are also mathematically tractable.

In Chapter 5 we discuss the steady state solution of the model formulated in Chapter 4. The more complex cases are investigated using simulation methods.

In Chapter 6, we simplify the previous model and use it to find maximum likelihood estimates of the 'infection' and 'recovery' rates using malaria survey data from the WHO Garki project.

We conclude in Chapter 7 with a summary of our results and some suggestions for possible future research.

CONTENTS

CHAPTER ONE

BACKGROUND TO MALARIA

1.1 Introduction

Malaria, probably one of man's most ancient diseases (Harrison, 1978, page 1), is caused by multiplication of parasitic protozoa of the genus Plasmodium, first in the liver, and then within the blood cells of its victim. Malaria infections are characterized by fever, enlargement of spleen, anaemia and a chronic relapsing course, and among children and non-immune adults they result in fatal complications in more than 10% of cases (Benenson, 1975, page 189).

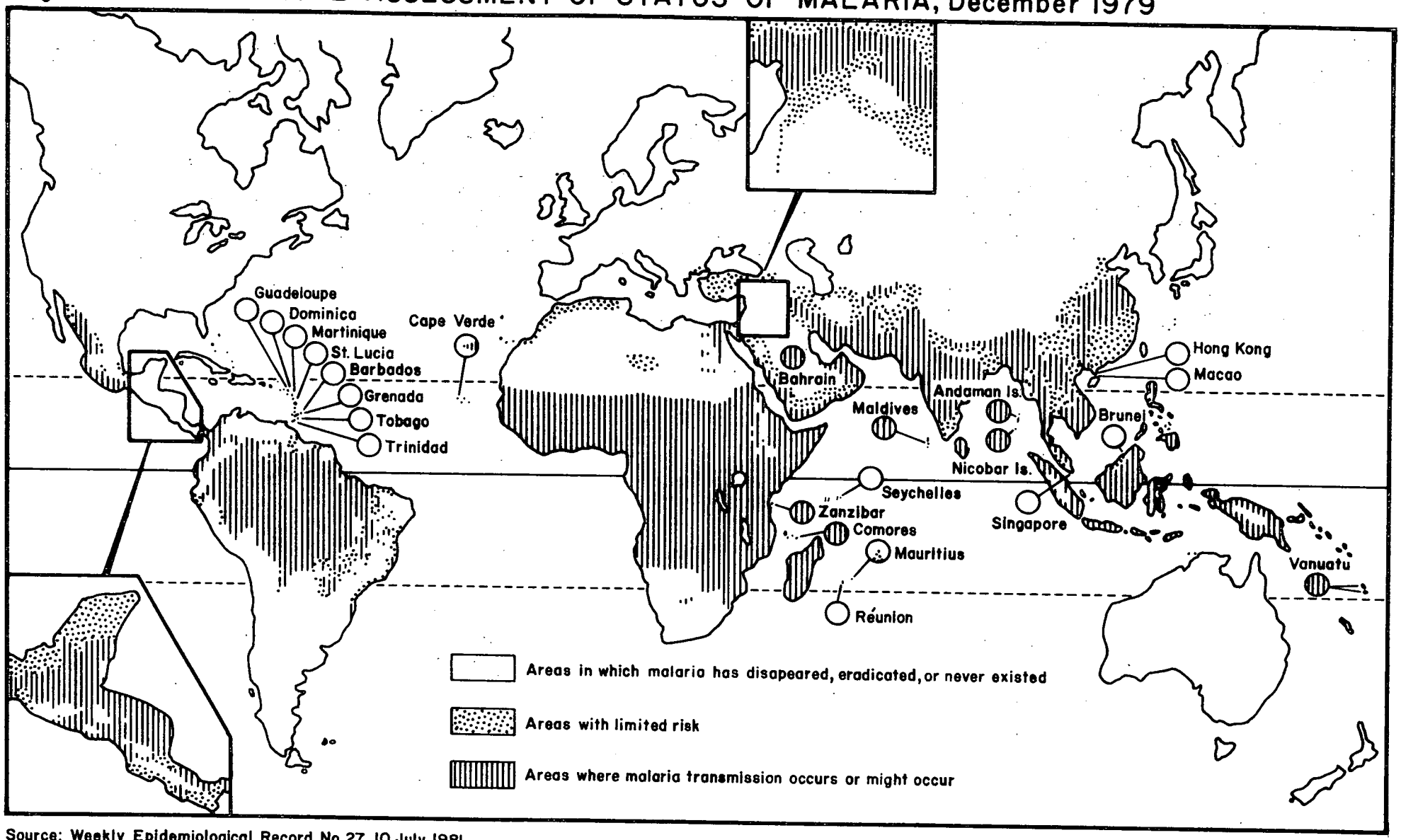
Although over fifty species of malaria parasites have been identified, only four are associated with man : Plasmodium (P.) falciparum, also known as malignant tertian, is the most lethal of the four; P. vivax, formerly known as tertian or benign tertian, seldom fatal; P. malariae, generally mild but stubborn quartan fever; P. ovale, the rarest of the four and found mostly in the West coast of Africa, is very much like P. vivax. It is generally accepted that man is the only main reservoir for these species (Benenson, 1975, page 190: Sambasivan, 1975). The various species of the parasites also differ in such details as their morphology, pathogenicity, incubation periods, relapses and duration of infections. In addition to being the most serious in severity of attacks, P. falciparum is the most prevalent species. Henceforth, and unless otherwise stated, we shall be concerned solely with infections due to this species alone.

The disease is transmitted naturally through the bites of infectious female mosquitoes of certain species of anopheles. Rare but direct transmission from man to man can however occur by either injection, or transfusion of blood of infected persons, or by the use of contaminated hypodermic syringes. Congenital transmission of the disease to the child may also occur when the mother has had bouts of malaria during the later months of pregnancy (Russell et al., 1963, page 455; Benenson, 1975, page 190).

Malaria is now confined largely to the tropical and subtropical areas of the world, although in the past it existed in the majority of the countries on the globe (Sambasivan, 1975; Harrison, 1978). It is estimated that over 400 million people still live in areas of high malaria infection risk (Bruce-Chwatt, 1980). Figure 1 shows the various areas where malaria poses serious health problems. The whole of tropical Africa is highly endemic with the disease, and it is a serious health hazard in parts of Asia, Central and South America. In tropical Africa, in particular, it is recognized as a major impediment to economic development because of the incapacitating and debilitating effects it produces in its victim. Moreover, instances of high prevalence of the disease usually coincides with periods of intensive agricultural activity.

The disease has thrived in these areas despite efforts through the World Health Organization (WHO), and local governments of these areas, to try and control or eradicate it completely (see Section 1.6 for control). There

Fig. I EPIDEMIOLOGICAL ASSESSMENT OF STATUS OF MALARIA, December 1979



Source: Weekly Epidemiological Record No. 27, 10 July 1981.

are various reasons for this lack of progress. First, there are the ecologically favourable conditions for mosquitoes (vectors) breeding coupled with the poor health services, inadequate material resources to wage a consistent war against the disease, and the lack of technical expertise in these regions. In addition, more recently a very serious problem has emerged of the parasites and mosquito vectors developing resistances to the drugs and residual insecticides, respectively, that have been traditionally used to combat the disease (Bruce-Chwatt, 1969, 1979; Sambasivan, 1975; Gilles, 1981). This, together with man's activities such as the setting of agricultural projects involving large bodies of water are all contributing to the resurgence of the disease in areas where it has been previously contained, for instance in Sri Lanka and some parts of India.

The disappearance of malaria in developed countries has been due to a combination of favourable factors, both natural and man-made. First, the low temperatures to be found in most of these countries limit the transmission of the disease in that P. falciparum requires a temperature of at least 20°C for its development in the mosquito (Sambasivan, 1975; Bruce-Chwatt, 1980) (see also Section 1.4). Secondly, the rapid socio-economic developments of the past few decades have reduced considerably the sources of the vector breeding. Finally, there are the extensive network of health services now available together with the active participation of a population motivated by a better awareness of their health needs. Nonetheless, the risk of importing cases is on the increase with the rise in volume

and speed of international travel. For instance, in Britain, 1,670 cases of P. falciparum malaria were recorded in 1980, of which seven were fatal. The mid-year figure for 1981 stood at 584. (Gilles, 1981).

In order to gain an understanding of the disease and appreciate its biological and ecological complexities, we shall provide, in the rest of this Chapter, brief outlines of its biological background, and factors influencing its transmission, both natural and man-made. The next section (1.2) carries a short historical note on early understanding of malaria. In section 1.3 we sketch the life-cycle of the malaria parasite, followed by the life-cycle of the malaria vector (section 1.4) and endemicity of malaria (section 1.5). In section 1.6 we outline some of the means of control available, both old and new, either through the human population (sub-section 1.6.1) or the vector population (sub-section 1.6.2). We conclude with a note on a WHO malaria study project carried out in Nigeria (sub-section 1.6.3) from 1970 to 1976.

1.2 A Historical Note

The name malaria comes from the Italian expression "mal' aria", meaning bad air. The French called it Paludisme from the latin word "Palus", which means marsh. These names are indicative of the views held in ancient times, which associated the disease with swamps, marshes and foul air. (Harrison, 1978, page 24).

Early developments of scientific knowledge of malaria

were dominated by two nationalities, first the Italians and, later, the British. For the Italians, this was probably because of the endemic malaria situation which had existed in Italy since ancient time, and as for the British, their vast colonies were almost all malaria infested. It was, however, the Frenchman, Laveran, who first discovered the parasites of malaria, while working in North Africa in 1880. He was probably influenced by an earlier discovery of his countryman, Kelsch, who observed some black pigments in the blood from corpses of malaria victims. Laveran's discovery took almost four years to gain acceptance as most researchers at the time believed that the disease was caused by some bacteria to be found in swamps, and not by parasites (ibid, page 11). This view of the disease was changed when, in 1884, an Italian, Camillo Golgi, reported observing and recording the asexual reproductive cycles of what he later called P.malariae, in human blood. Soon after he also discovered those of P. vivax and malignant malaria of the type originally identified by Laveran (ibid, page 15).

According to Harrison, the suspicion that mosquitoes were somehow connected with malaria was ancient and fairly common among people in malaria infested areas. However, it was Patrick Manson who, in 1894, first conceived the idea that mosquitoes were actual carriers of malaria. It is conceivable that Manson's own discovery in China, during the 1870's, that mosquitoes were carriers of filarial worms, gave him the cue. But he thought people contracted the disease through contact with water previously contaminated by dead infected mosquitoes. Another Italian, Amico Bignami

countered this with an alternative view that malaria was passed to humans by mosquitoes during feeding. It is interesting to note that between them, Bignami and Manson had in fact solved the riddle of malaria transmission as we know today, from man to mosquito (Manson) and from mosquito to man (Bignami), but neither thought the other was right.

It was left to Ronald Ross who was himself introduced to malaria in 1894 by Manson to demonstrate experimentally, using birds, the transmission of malaria by mosquito in the way suggested by Bignami. Soon after, Bignami himself and a fellow countryman, Grassi, showed that the disease was similarly transmitted to humans by mosquitoes. This was followed in 1898 by the discovery of Grassi that the Anopheles mosquito was the sole carrier of malaria.

It must be pointed out that some of these discoveries were fraught with fierce controversies and credits for them were also often in dispute. In addition, it must be emphasized that those mentioned here are but a few of the many who pioneered scientific research in and made significant contributions to our understanding of malaria. Among these is the Canadian, MacCallum who, in 1897, was the first to observe the process of sexual conjugation of malaria parasites, the American, Koch, who demonstrated the existence of the chromatin substance in the flagella of the parasite and the Frenchman, Maillot who pioneered the development of successful quinine therapy to the disease. The book by Harrison (1978) provides a lucid account of the history of the disease.

1.3 Life-cycle of the malaria parasite

The malaria parasite undergoes a complicated cycle of development that requires two hosts. The definitive host or the mosquito vector is where the parasite achieves sexual maturity, and the process is known as sporogony. The asexual multiplication of the parasite or schizogony, takes place in man, the intermediate host. Schizogony has two distinct phases of development; the tissue phase, called the pre-erythrocytic phase, which occurs in the (parenchyma) cells of the liver, and the erythrocytic phase which takes place in the red blood cells in the main blood streams.

To be able to understand the various stages of the development of the disease, we shall sketch below only the general outline of the life-cycle of the parasite. Figure 2 is a schematic diagram of the cycle. An excellent and concise book on the subject is that by Bruce-Chwatt (1980) from which much of what follows is drawn.

(a) The Human Host

A man is infected by an inoculation of sporozoites - the final product of the process of sporogony - when bitten by an infectious female anopheline mosquito. These spindle-shaped parasite forms circulate in the peripheral blood for about 30 minutes, during which some are destroyed by phagocytosis. Those that survive enter and lodge in the parenchyma cells of the liver, to initiate the pre-erythrocytic stages of their development. Within the parasitized liver cells, the nuclear material of the Plasmodium divides repeatedly to form between 1,000 to 40,000 daughter forms

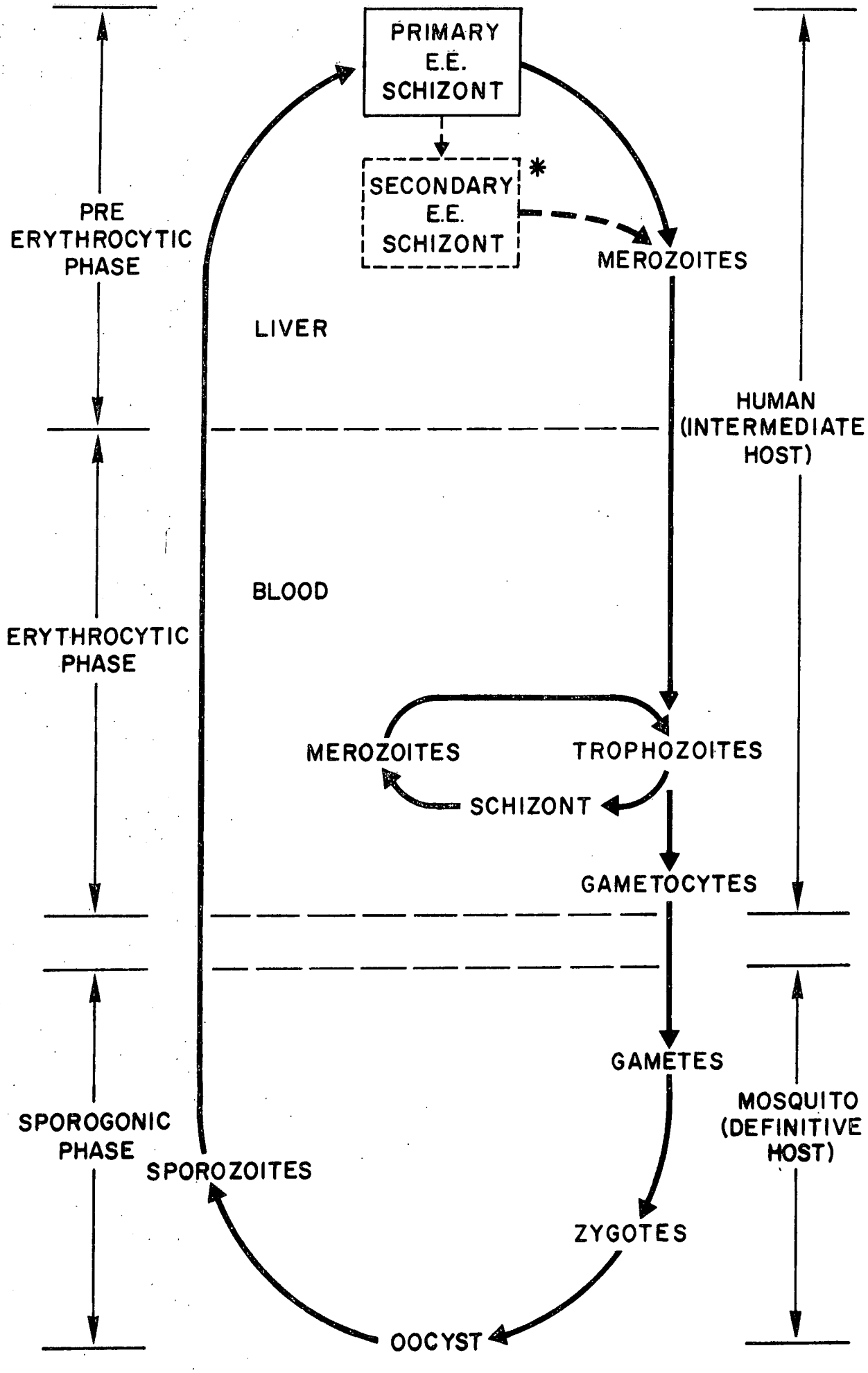


Fig 2 LIFE CYCLE OF THE MALARIA PARASITES – P. FALCIPERUM

* *Note in P. ovale and P. Malaria the primary E.E phase is followed by secondary E.E. as explained in text*

called merozoites. After about 5½ to 7 days of development, the infected cells burst to release thousands of these merozoites into the surrounding tissue and blood-stream. Most of these merozoites, now known as trophozoites, then invade red blood cells to begin the erythrocytic phase; others get phagocytized.

Incidentally, in the case of P. vivax and P. ovale, pre-erythrocytic stages are usually followed by a secondary exo-erythrocytic (E.E.) phase, during which some of the merozoites re-infect fresh liver cells instead, thus repeating the process (see Figure 2). However, more recently a different explanation of this late liver schizogony has emerged. Some malariologists now believe that it is caused by late-developing sporozoites, rather than by re-invasions of the liver cells by merozoites (Bruce-Chwatt, 1980). Thus, the idea that relapses are caused by secondary exo-erythrocytic phase is now a matter of some contention. Moreover, it is not yet clear what causes this condition of 'dormancy' to occur in one species and not in others.

Each trophozoite that invades and settles in a red blood cell grows and multiplies asexually to form a developing schizont. After about 48 hours the red blood cells containing the matured schizonts rupture, each releasing up to 24 merozoites plus some organic debris into the blood plasma. Some of these merozoites are phagocytized while others attack fresh red blood cells, initiating the production of another generation of merozoites. The cycle is repeated until a very heavy blood infection results which, if unchecked by either drugs or immune mechanism (see sub-section 1.5.1),

may mean the death of the host. It has been suggested that in non-fatal cases, and in the absence of re-infection and intervening treatment, the duration of infection in falcipārum malaria seldom exceeds 1 year (Russel et al., 1963, page 40). The maximum number of parasites in a host is usually found about 8 days after parasites are detected microscopically (Belding, 1965).

Clinical symptoms of the disease often appear after 2 to 3 days of erythrocytic phase depending on the resistances of the host. At this stage the density of parasites in the body is such that paroxysms of fever ensue. The host becomes shivering cold with body temperature rising up to 40°C or more. After several hours he begins to sweat profusely. This is followed by some relief and a feeling of weakness. Fever then resurges about every day accompanied by severe headaches, frequent vomiting and anaemia. The latter is due to the destruction of the red blood cells by the process of schizogony. If untreated the disease may become severe and debilitating. In the case of P.falciparum, infected red cells get trapped in tissues and narrow vessels of the inner organs, thus blocking and slowing down the flow of blood. If this occurs in the brain, the host goes into a coma and death soon follows. In general, the incubation period, which extends from the time of infection to that of first appearance of clinical symptoms, varies from 9 to 14 days, with an average of 12 days (Russel et al., 1963, Ch.12).

Several erythrocytic cycles later, gametogony sets in. After invading fresh red blood cells, some merozoites do not divide; instead, they grow into male and female gametocytes.

Little is yet known as to what actually triggers their production or the precise point at which gametogony begins. They appear to arise spontaneously from asexual division. In the case of P. falciparum, the sexual forms appear about 10 days after the onset of the erythrocytic phase. Once formed, gametocytes do not undergo any further development in the host except circulate in the peripheral blood, where they can be sucked into the mid-gut of a mosquito when it takes blood meal from the human, thus initiating the sporogonic phase of the parasite development.

The number of gametocytes produced is variable and is usually high in falciparum infections, much more so than in any infections by the other species. As to their duration of life, the studies of Jeffrey et al. (1956) and, more recently, Smalley et al. (1977) estimate it to be about 20 days. Their finding is at variance with, for instance, that of Hawking et al. (1971) which says gametocytes are short-lived, or that of Russell et al. (1963) which suggests the figure of 120 days. Smalley et al. also observed that gametocytes are removed in the spleen, which is to be expected since one major function of the spleen is the removal of debris and parasitic invaders.

(b) The vector

Once inside the mosquito each male gametocyte produces up to 8 thread-like micro-gametes while the female form develops into a macro-gamete. Fertilization occurs when the nuclei of a micro-gamete and a macro-gamete fuse to produce a motile, elongate cell called ookinete. The ookinete develops to form an oocyst which in turn divides repeatedly

to form sporozoites. When a mature oocyst ruptures it releases thousands of spindle-shaped sporozoites into the body cavity and a large number of them find their way into the salivary glands of the mosquito. When taking its next blood meal, the mosquito injects sporozoite-burdened saliva into its victim (man), thus completing the cycle that perpetuates the Plasmodium indefinitely unless the cycle is interrupted somehow. The entire sporogonic process takes from 10 to 14 days.

1.4 Malaria Vector

There are more than 200 recorded species of anopheles mosquitoes of which about 50 are considered responsible for the transmission of malaria parasites in man (Russell et al., 1963, page 269). Some species of anopheles are also vectors of filariasis and certain virus, for instance Plasmodium aedes in Yellow fever (Service, 1980, page 22). We give below a brief description of their life-cycle.

The full life-cycle of a mosquito involves 4 stages of which the first three are aquatic : the egg, larvae, pupa and the adult stages respectively. After the eggs are laid on water, under favourable conditions, they hatch into larva within 2 days. The larvae stage lasts about 5 to 7 days, and the pupa stage, 2 days. Temperature and species are the main determinants of the durations of these aquatic stages. In all, it takes from 9 to 11 days to complete the life-cycle. After that the adult may live, from a few days to a few months, depending on the temperature and humidity.

The adult is particularly vulnerable to dry weather.

However, it can survive through dry as well as cold weather, simply by adjusting its pattern of life and adopting a depressed state of physiological activities (Sambasivan, 1975). Of crucial importance perhaps is the longevity of the adult female vector since it is the only one that bites blooded animals to obtain blood for the development of its ovaries. Even under the most favourable conditions, it cannot transmit malaria until it has survived the period of sporogony. Once infected a female vector remains infected and a transmitter of the disease for the rest of its adult life.

The significance of temperature and humidity in connection with the transmission of malaria also extends to the development of the parasites of the disease inside the mosquito. In general, P. falciparum requires a minimum temperature of about 20°C to develop in the mosquito. The parasites cease to develop below the temperature of 16°C. The ideal temperature is considered to be from about 20°C to 30°C and a mean relative humidity of at least 60%. These two factors, namely temperature and humidity, combine to link malaria with rainfall, particularly in the tropics and sub-tropical areas, a point which we shall exploit in subsequent chapters.

1.5 Endemicity of Malaria

Since we shall be primarily concerned with endemic rather than epidemic malaria, we propose to describe briefly some of the main characteristics of endemic malaria and some of the factors influencing endemicity. We begin by stating the epidemiological distinction between endemic and epidemic malaria.

When malaria is said to be endemic in a community, it is usually understood to mean the habitual presence of malaria within the community due to natural local transmission. It is epidemic when its occurrence in the community is in excess of normal expectancy. Both cases are a result of interactions of various factors affecting the bionomics of man, mosquito and the parasite in the local environment.

Endemicity is further classified according to the degrees of incidence and intensity of cases of the disease. The classification ranges from hypo-endemicity, in which transmission is low and the effect of malaria is unimportant, to holo-endemicity, where transmission is perennial and of high intensity, resulting in marked level of immunity responses in all age groups, especially in adults.

1.5.1 The Immune Mechanism

An important aspect of endemic malaria is clearly that of immunity. Unlike in the case of other diseases such as yellow fever, immunity in malaria does not bestow complete protection from the disease upon a person. It serves rather as an active defence mechanism to limit the multiplication of Plasmodium, modify its effects on the body and generally assist in repair of tissues damaged by the disease (Russell et al., 1963, page 425). Immunity is either natural or acquired. We proceed to give brief descriptions of each type.

(a) Natural Immunity

Natural immunity is independent of previous infection. It is either passed to the child from the mother, through the placenta, or it is of genetic origin. It is known to

act on both the asexual and the sexual erythrocytic stages of the parasite (Manson's Tropical Diseases, 1982, page 45). This presence of natural immunity in children in endemic areas probably explains why such children enjoy protection from the severity of the disease until they are about 3 months old when the immune mechanism acquired from their mothers begins to weaken.

One genetic characteristic known to mitigate the severity of infection in man is the presence of haemoglobin S (sickle-cell trait) in the blood. This abnormal haemoglobin, common among communities in East and West Africa, has been observed to protect the carriers against malaria; its presence in the blood apparently does not favour the growth and development of the parasites in the red blood cells (ibid, page 427). However, it gives rise to a condition called a crisis, a painfull state during which the patient is gravely incapacitated and cannot rest unless given a strong pain killer. Although it has been postulated that hereditary factors other than haemoglobin S, such as haemoglobin C and Glucose-6-phosphate dehydrogenase (G-6-PD) deficiency, haemoglobin E, and thalessamia protect against lethal effects of malaria, the only convincing evidence to date concerns sickle cell haemoglobin and G-6-PD deficiency (Lucas et al., 1981, page 179).

(b) Acquired Immunity

By far the most important defence mechanism in man is acquired immunity. Its main function is to accelerate the phagocytosis of merozoites and debris in the blood and also reduce the average number of merozoites produced by a

mature schizont (Garnham, 1966). Acquired immunity is built up gradually in the body as a result of a series of infections over time. According to Garnham, it is produced only during the asexual erythrocytic stages of the malaria parasites. So far as is known, neither the liver stages nor gametogony produce it, nor are they affected by it.

There is evidence that acquired immunity is associated with significant increase of gamma-globulin and various types of antibodies in the blood (Collin and Skinner, 1964 ; Bruce-Chwatt, 1965). Field studies made largely in Africa over the past two decades have revealed a positive correlation between the rising concentration of serum gamma-globulin and the acquisition of clinical immunity to P.falciparum (McGregor, 1974; Molineaux et al., 1980, Chapter 6); see Chapter 6, section 6.5.

The situation in an endemic area may be as follows. From the age of 6 months to about 2 years, children become very vulnerable to acute and fatal attacks of malaria as their natural immunity (acquired from mother) wears out with age. Thereafter, those who survive, build up acquired immunity as a result of re-infections (and relapses in the case of P.malariae and P.vivax). At first it serves to modify the severity of clinical symptoms, then, later, reduce parasitamia (Dietz et al., 1974).

One important feature of acquired immunity is that it enables malaria to be a self-limiting disease. In the absence of fresh transmission, even without specific treatment, it tends to clear away an infection after some time. It has been suggested, as mentioned earlier, that the life-span of most untreated, non-fatal infections of

P.falciparum seldom exceeds 1 year (Russell et al., 1963). However, on its own, once established, immunity may be lost over a period of 2 years or more after the infection has been cured and transmission eliminated (Sambasivan, 1975). Thus persons who have left an endemic area for more than three years will have lost their acquired immunity and may suffer a few severe attacks on returning to the area.

Finally, it is worth mentioning that acquired immunity is generally strain-specific. That is, while it may serve to protect a person from the severity of infection of a local strain, it offers no similar protection against a strain from another area (Sambasivan, 1975). Furthermore, as mentioned before, immune response to P. falciparum gametocytes is uncertain. According to Smalley et al., (1977), although some form of antibodies against gametocytes are produced, these are either not as effective as those for merozoites, or not produced by every gametocyte carrier.

1.5.2 Human Activities

A number of man's habits affect the spread and perpetuation of the endemicity of the disease. For instance, persons whose occupations keep them out-doors at night run a great risk of being bitten by infectious mosquitoes. Or, if they happen to be infected, they may infect more mosquitoes which in turn may infect other susceptible members of the community. Ironically, some of the activities of man meant for economic development, such as the construction of dams, the irrigation systems and other engineering works involving large bodies of water, usually boost the malaria risk of an area unless adequate measures are taken to

prevent the creation of new mosquito breeding places (see next section on control measures). Finally, the movements of small or large groups of persons in search of work, grazing grounds or because of civil disturbances or disasters, often occur in most of these malarious areas (see Fig. 1). Consequently, new strains are frequently traded around, thus perpetuating the disease to the detriment of the efforts of those engaged in either controlling or eradicating the disease.

It must be emphasized that what has been said of the disease so far is only a very simple form of an otherwise complex biological process which still has plenty of un-answered questions.

1.6 The Control of Malaria

A wide range of measures aimed at either controlling or eradicating malaria have been developed over the years. However, the contemporary history of the fight against the disease is more or less that of retreat. The international effort to rid the world of malaria has been spearheaded by the World Health Organization (WHO), and malaria has been for some time now put on the list of "diseases under surveillance". In 1955 and 1956 the 8th and 9th World Health Assemblies adopted a policy of malaria eradication. In 1969 they revised their policy down to that of eradication plus control. A year before, Sri Lanka, one of the areas in which malaria was considered wiped out, experienced an epidemic of P. vivax malaria affecting more than 1.5 million people (Bruce-Chwatt, 1975; Harrison, 1978; Gilles, 1981). The idea of total eradication of the

disease particularly in tropical Africa is now considered unattainable. Emphasis is on containment and control aimed at reducing the morbidity and mortality of the disease in these areas. Some of the socio-economic and ecological reasons for this change of strategy have already been enumerated in our introductory remarks. We shall outline below some of the measures being used to control the disease. The measures may be grouped under two headings; namely those for the protection of susceptible individuals and those directed against the vector as strategies to protect entire communities.

1.6.1 Susceptible Individuals

People who move into endemic malaria zones from areas free of the disease do not have resistance to the lethal effect of the disease and therefore need to take extra precautions. The risk also extends to individuals who normally live in these zones but have been away in malaria-free areas for a long time. The regular intake of anti-malaria drugs such as chloroquine, daraprim (taken once weekly) and paludrine (daily) minimizes this risk. Other measures include the siting of dwellings away from potential breeding places of the vector, screening of living quarters and the use of bed nets, protective clothing and mosquito repellents at night to prevent mosquito bites, and the use of insecticides in aerosol dispensers.

There are a number of preventive measures being researched of which the development of vaccines against malaria is one (Gilles, 1981). An overwhelming difficulty with

the development of malaria vaccines is that antigens on which these vaccines are based are often stage specific. Thus, a vaccine based on antigens from the liver stages may have no effect on the erythrocytic phase of the disease (Cohen, 1979) and vice versa.

The most promising form of malaria vaccine is that based on antigens from the erythrocytic phase of the parasites. The possibility of the production of these antigens for mass vaccination has been enhanced by the development of a technique of continuous cultivation of erythrocytic forms of P. falciparum (Trager and Jensen, 1976). However, there are still some problems to be surmounted.

Using current techniques, it is difficult to exclude pathogens from materials prepared for the vaccine and the risk of contamination with blood group substances is high. Much more serious perhaps is the situation that immune response of a given population is likely to be influenced by such factors as genetic differences, blood group specificity and haemoglobin combinations (Piazza et al., 1972). At present the vaccine is being tried on animals. Cohen (1979) has warned of its use being extrapolated to man without careful evaluation and thorough preparations.

The treatment of acute and chronic cases of the disease may be affected by the use of drugs such as quinine, amodiaquine, chloroquine, quinine sulphate or dihydrochloride (Benenson, 1975, page 194). In parts of South America, South-east Asia and the western Pacific regions, where P. falciparum infections are chloroquine-resistant, the use of quinine followed by a regime of pyrimethamine

is suggested (Gilles, 1981; Benenson, 1975, page 194). A recent drug, mefloquine, found to be effective against multidrug-resistant strains of P. falciparum, has yet to be produced on a commercial scale. Up-to-date descriptions of treatments of P. falciparum infection have been published by Hall (1976) and Bruce-Chwatt et al. (1981).

1.6.2 The Vector

Control measures against the vectors are directed at the vector breeding and their larvae and adult forms.

(a) Prevention of breeding

Various methods to prevent vector breeding include drainage and filling breeding places which result in reduction of anopheline breeding habitats. However, usually a good knowledge of the breeding preferences of the local vector is essential for selecting the method of control. For example, for Anopheles (A.) umbrosus in Western Malaysia and A. leucosphysus in Sabah (Malaysia) that prefer shaded breeding places, clearance of overhanging vegetation has been used with considerable success (Sambasivan, 1975). Another method sometimes used is to colonize mosquito breeding places with predatory fish such as guppies.

(b) Larvicide

Measures to control the vectors with larvicides are also common. They involve, for instance, the use of mineral oils to form films over water surfaces which suffocate and ultimately kill any larva or pupa present in the water. Other substances include Paris green, DDT and some more recently

developed organo-phosphorus compounds such as temephos and fenthion. But DDT has long term adverse effect on aquatic life and is less used in this fashion than as a residual insecticide (see following subsection). On the whole, the effectiveness of anti-larval method varies with the particular vector species involved.

(c) Use of Residual Insecticides

Probably the single most important control measure directed at the adult vector is the use, once or twice a year, of residual insecticides. Usually in liquid form, these insecticides are sprayed on walls of dwellings where their residuals remain for a long time. Mosquitoes get poisoned and eventually die when they come to rest on these walls before or after meals. DDT has been by far the most widely used and the most popular. It is easy to use and transport. It also has prolonged residual effect and is the least expensive of all insecticides. It has been successfully used in the past in eradicating the disease in a number of countries in Central America and parts of the Carribean (e.g. Cuba). (Sambasivan, 1975).

In areas where DDT has either been discouraged because of side effects or abandoned because of vector resistance to it, organo-phosphorus based compounds such as Malathion and Dichlorvos (DDVP) are being tried. In particular, Malathion has low toxicity to man but it is more expensive than DDT. Other replacements are the more recently developed carbonate compounds such as propoxur (OMS33). Propoxur is a very efficient residual insecticide. For instance, it can kill mosquitoes 10 to 30 metres outside a house sprayed

with it. However, it has the disadvantage that it is 8 times more expensive than DDT and this puts propoxur beyond the reach of most malaria-ridden developing countries.

(d) Other Methods

These other methods are mainly aimed at inhibiting the evolution of resistance in mosquitoes to insecticides, and they are still at their development stages. The goal is to extend the effectiveness of existing insecticides. This is motivated by the fact that the costs of developing and purchasing new insecticides are becoming prohibitive. Moreover, resistance by a vector to a particular insecticide may also extend to a new insecticide as in certain DDT resistance cases and synthetic pyrethroids (Chadwick et al., 1977, Prasittisuk and Busvine, 1977).

It has been noted by Muir (1975, 1977), Comins (1977) and others (Curtis et al., 1978) that migration of insects in populations untreated with a given insecticide would tend to delay the evolution of resistance in a neighbouring colony of treated insects. It has been suggested therefore that insecticides should be applied in alternating sectors of a "grid" in which adjacent sectors are sprayed with, for instance, chemically unrelated insecticides. One other proposal suggested by Curtis et al. (1978), is that of reinforcement of natural immigration by the release of artificially reared male mosquitoes that are susceptible to a specific insecticide. These are supposed to re-introduce susceptibility of genes into the progeny of the resistant population. With the technique of rearing mosquitoes on a

large scale becoming well established (Singh et al. 1974, Dame et al., 1974), the latter method appears promising. However, the idea of increasing the population of man-biting insects in a community may not at first be particularly appealing.

1.6.3 The WHO Malaria Surveys in Garki

As part of its global effort to control malaria, the WHO initiated in 1969, what is now known as the Garki Project. The project, which was multi-disciplinary in scope, was based on a series of longitudinal field surveys of the human population, together with mosquito and meteorological surveys carried out in Garki, Kano State, Nigeria, in conjunction with the Government of the Federal Republic of Nigeria.

The project had three specific objectives (Molineaux et al., 1980, page 21) -

(1) to study the epidemiology of malaria, which involved the measurement and study of entomological, parasitalogical and serological variables and their relationships. It also included the measurements of meteorological, demographic and clinical variables.

(2) to measure the effects of such control measures as the use of residual insecticides with and without mass drug administration.

(3) to construct and test a mathematical model that would simulate the transmission of the disease under various control strategies.

A comprehensive summary of the results of the studies carried out so far by various members of the project team, in pursuit

of the above objectives, has recently been published as a WHO monograph (Molineaux and Gramiccia, 1980).

The demographic-parasitalogical surveys were aimed at covering the entire population of sixteen selected villages in the area, and were carried out in three phases. In the first phase, surveys were conducted every ten weeks from the end of the wet season in November 1970 to the end of the dry season in May 1972. Throughout that period no attempts were made at treatment (except probably in the most serious cases) or control of malaria. Consequently, these surveys are called the baseline surveys. The second phase known as the intervention phase, in which surveys were also conducted every ten weeks, commenced from April 1972 to October 1973, during which certain control strategies were applied to some of the villages while others were left untreated as control. The third and last phase, or post-intervention phase, extended from November 1973 to the end of the project life, February 1976. During this period selective drug administration was carried out, particularly in villages covered by mass drug administration in the second phase. The aim was to protect these villages whose populations' immune responses had apparently been weakened from lethal malaria attacks. Throughout the project infants were surveyed twice as many times as the adults, that is, about every five weeks.

We shall subsequently (Chapter 6) be concerned only with the data from the baseline phase, and in particular the parasitalogical data. These data, (as well as those in the other phases) were obtained by the collection and examination, at each survey and for each person, of a thick blood film, linked

by an identification code number of the person. The blood film was examined by means of a microscope, for 200 fields, and the person was then classified, at that particular survey, as positive for P. falciparum asexual forms, P.falciparum gametocytes, P. malariae or P. ovale respectively, if any of these were observed in any of the 200 fields, and the number of positive fields recorded. Otherwise he was classified as negative, that is, free from parasitamia.

As regards those surveys for gathering sero-immunological data, these were carried out in selected villages twice a year, in conjunction with the demographic-parasitalogical surveys, one in the dry season and the other in the wet season. The surveys involved the collection of blood samples from fingerpricks into heparinized caraway tubes and filter papers on which several tests were then performed. Among the tests were those to determine quantitatively the levels of immunoglobulins G and M, using techniques described in the paper by Mancini et al. (1965).

The information gathered in the Garki Project is believed to provide a better basis than was previously available, for the quantitative study of malaria and the planning of malaria control. Before the project was undertaken all malaria surveys on man were, with a few exceptions (Krafsur and Armstrong, 1978), cross-sectional prevalence surveys (Russell et al., 1963). The Garki data do provide the opportunity to obtain good quantitative assessments of some of the key factors governing malaria transmission such as age and seasonality. In Chapter 6, we shall use some of the data to estimate such parameters as human rates of infection and recovery rates by age-group and season.

CHAPTER TWO

REVIEW OF QUANTITATIVE STUDIES OF MALARIA

2.1 Introduction

Early attempts to apply mathematical techniques to the transmission and control of malaria are attributed to Ross (1911, 1915, 1916, 1917). Much of later studies were extensions of some of his original ideas. Macdonald (1950, (a,b), 1953, 1955, 1957, 1965) considered various deterministic aspects of the spread of endemic malaria the essence of which could be traced back to Ross's pioneering work. Other workers include Martini (1921), Lotka (1923), and Moshkovskii (1950, 1967). Lotka, in particular, studied extensively the various equations proposed by Ross and illustrated his solutions of the equations by numerical examples. Waite (1910) considered the relationship between the number of mosquitoes in a locality and the malaria rate, also based on Ross's ideas. Others who have devoted some attention to malaria include Armitage (1953) and Bailey (1975).

Kermack and McKendrick (1927) have discussed a simple deterministic model of the transmission of malaria and obtained an approximate result for the threshold of an epidemic; that is, some level of the disease prevalence below which it dies out but above which it maintains itself. A stochastic analogue of the deterministic threshold has been derived by Bartlett (1964, 1966). Oyelese (1970) and Rao et al.(1974a,b) have also derived various forms of the threshold theorem also using stochastic models.

Recent developments on the subject are due to Dietz and

his colleagues at WHO (1974) whose model is specifically based on the Garki project described in the last Chapter (section 1.6). More recently, Aron and May (1980) have reviewed the basic model of Macdonald. They have also modified it slightly by incorporating such aspects of the disease as latency and immunity. An earlier review of the subject is contained in the book of Bailey (1975, Chapter 17).

Attempts have also been made to model the parasite population in the intermediate host. Elderkin et al.,(1977) have modelled the parasite population and the acquisition of immunity in which both the parasite population and immune response vary continuously with age. Before that, Marcus (1970) has used the method of branching processes to model the cyclic development of the parasites inside a mammalian host.

In the next four sections we outline briefly the essential aspects of some of the above-mentioned models. We begin with the basic deterministic model of Ross.

2.2 The Basic Model of Ross with Extension

Ross's basic model is defined as a set of differential equations. With t representing the time variable, the equations are :

$$\frac{dx}{dt} = b'f'u(1-x) - \gamma x \qquad (1)$$

$$\frac{du}{dt} = b'fx(a-u) - \nu'u \qquad (2)$$

where

x : proportion of human population infected;

u : the density of infected mosquitoes per head of human population;

a : the overall female mosquito density per person;

f : proportion of infected persons who are also infectious;

f' : proportion of infected mosquitoes that are also infectious;

γ : human recovery rate;

b' : mosquito man-biting rate;

ν' : mosquito death rate

By using the proportion infected in the human population, x, and the density of infected mosquitoes, u, as the dynamic variables, it is assumed that the populations involved are, each, approximately constant. The first equation describes changes in the proportion of persons infected. That is, new infections are acquired at a rate that depends on the mosquito man-biting rate, b', the density of infectious mosquitoes, f'u, and the proportion of persons not infected, (1-x), from which is subtracted the proportion of infected persons who have recovered, γx. Equation (2) describes changes in the density of infected mosquitoes. These depend on the proportion of infectious persons, fx, the man-biting rate, b', the density of uninfected mosquitoes, (a-u), and the loss term, $\nu'u$, which results from the death

of infected mosquitoes. Throughout the sequel, by either a mosquito or a vector, we shall mean the female anopheles mosquito.

To achieve the simplicity in the preceding formulation, a number of key assumptions are made, in addition to that of constant human and vector populations. For the human populations, the possible effects of migration are ignored and no new susceptibles are allowed. It is also assumed that the death rate of humans is negligible relative to their recovery rate. As for the mosquitoes, once infected they are assumed to remain infected for the rest of their lives. Although the mortality rate among infected mosquitoes is believed to be higher than that among healthy ones (Anderson and May, 1979), they are assumed to be the same in the model. Furthermore, an uninfected mosquito gets infected whenever it bites an infectious host. It has been suggested, however, that the infectiousness of a disease to a mosquito depends partly on the age of the infective, especially in endemic areas. According to Sambasivan (1975), gametocytes from long-standing infectives are less infectious to the vector than those of younger infectives. This is probably why, in endemic areas, children show greater gametocyte densities than older members of the community, thus serving as the main reservoir of infection. Also neglected in the model is the length of the sporogonic cycle in the vector even though it is comparable with the life-span of the vector. Nor does the model take account of the various stages of development of the parasites in the human host such as the latency period. Thus the model ignores

completely the host-parasite relationship of the disease. Despite these short-comings, the model reflects the basic features of the interactions between the populations of human hosts and mosquitoes.

Equilibrium states of the disease may be found by setting $\frac{dx}{dt} = \frac{du}{dt} = o$ in equations (1) and (2), and solving for x and u. This results in two sets of solutions.

$$(i) \qquad x = u = o \tag{3}$$

and

$$(ii) \qquad x = \frac{ab'^2 ff' - \gamma\nu'}{b'f\,(\gamma + ab'f')} \tag{4a}$$

$$u = \frac{ab'^2 ff' - \gamma\nu'}{b'f'(\nu' + b'f)} \tag{4b}$$

Lotka (1923b) carried out studies of approximate time dependent solutions of equations (1) and (2), near these points of equilibrium, with ν' replaced by μ'. the birth rate of mosquitoes. This substitution was made by Ross (Bailey, 1975 p.315) who assumed that the birth-and death-rates of mosquitoes exactly balanced each other. It is unclear as to what this substitution was meant to achieve. In his investigation, Lotka found that for a given set of parametric values in the expression,

$$\frac{ab'^2 ff'}{\gamma\mu'} \tag{5}$$

the solution to these equations tends only to one of the two equilibrium solutions, depending on whether (5) is less than or equal to unity, or not. More specifically, the solution tends to (4) if expression (5) is greater than unity, otherwise the trivial solution is stable. This has been interpreted to imply that, if (5) does not exceed 1, then

a few malaria cases introduced into a malaria-free community will not result in any epidemic or endemic situation, and the disease soon dies out. Otherwise such an introduction will culminate in the disease reaching endemic proportions given by (4) (Bailey, 1975, p.316).

2.2.1 The Basic Reproduction Rate

Expression (5) has given rise to the concept of 'basic reproduction rate' of Macdonald, (Macdonald, 1952, 1957), which is conventionally denoted in the literature by Zo. If we replace μ' in (5) by ν', then the basic reproduction rate is given as

$$Zo = \frac{ab'^2ff'}{\gamma\nu'} \qquad (6)$$

This is usually interpreted as the average number of secondary infections contracted from a single infected person in a large population of susceptibles. If Zo is less than unity, the disease soon dies out, otherwise it will persist to assume endemic proportions ultimately.

An intuitive interpretation of Zo given by Bailey (1975, p.317), is as follows. For a single primary case with a recovery rate of γ, the average infection time is $1/\gamma$, during which he will have infected $ab'f/\gamma$ mosquitoes. Each of these mosquitoes lives an average of time $\frac{1}{\nu'}$, during which it will have administered a total of $b'f'/\nu'$ infectious bites, largely on susceptible humans. The total number of secondary cases is provided by the product $(ab'f/\gamma)(b'f'/\nu')$.

Recently Aron and May (1980) have derived Zo using a geometric phase-plane analysis. The horizontal axis corresponds

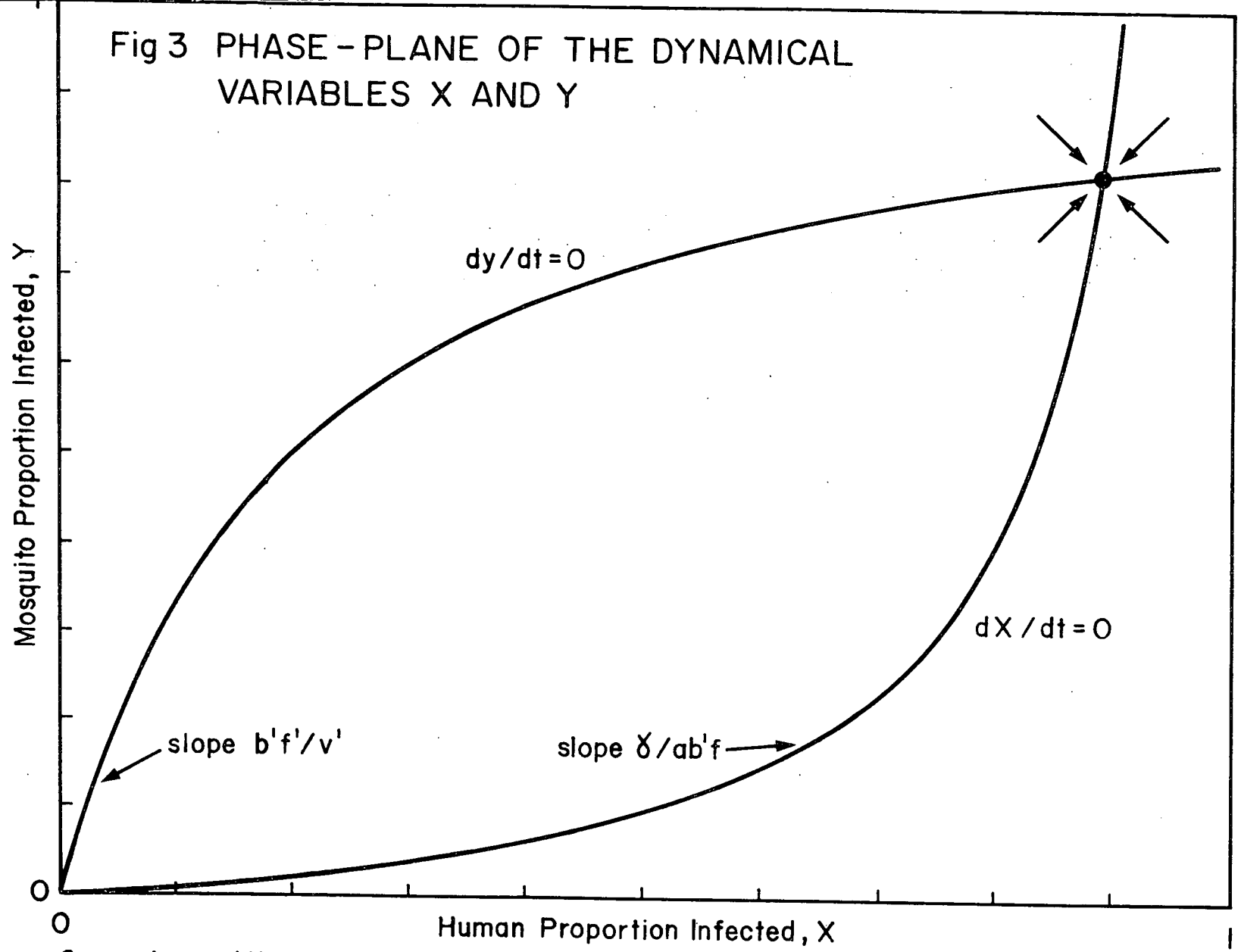
to x, the proportion of infected persons, and the vertical axis represents the proportion of infected mosquitoes, u/a, which they denote by y. The area of the graph is split into 4 domains by 2 isoclines corresponding to $\frac{dx}{dt} = o$ and $\frac{dy}{dt} = o$. The intersections of these isoclines are the equilibrium points. Equation (6) is obtained by dividing the initial slope of the y-isocline ($b'f'/\nu'$) by that of the x-isocline ($\gamma/ab'f$). Their result is reproduced in figure 3.

Their analysis suggests that small changes in either the man-biting rate, b', or in the mosquito density, a, are more likely to result in significant changes in the proportion of persons infected when b'/ν' is small than when it is large. This, they considered, is a vindication of Macdonald's argument in support of using b'/ν' as an index of stability. Thus in areas where mosquitoes biteman relatively often (larger b') and have long expected life (small ν'), b'/ν' is high and malaria tends to be endemic or 'stable'. Epidemic outbreaks occur in areas where mosquitoes bite less and have shorter life-span (b'/ν' small) ('unstable' malaria).

The parameter f' is in essence the latency parameter for the mosquito. Let the incubation period of the parasite within the mosquito (sporogonic period) be assumed constant and denoted by τ. Since an infected mosquito becomes infectious only after the completion of the sporogonic phase, f' may be interpreted as the probability that it is still alive by then, and approximately

$$f' = \exp(-\nu'\tau), \qquad (7)$$

Fig 3 PHASE-PLANE OF THE DYNAMICAL VARIABLES X AND Y



Source: Aron and May, 1980

so that

$$Zo = \frac{ab'^2 f \exp(-\nu'\tau)}{\gamma\nu'} \qquad (8)$$

It is from (8) that Macdonald (1957) argued that, as control measures directed at the mosquito vectors, 'insecticides are more effective than larvicides'. The larval component enters Zo linearly through a, so that the reduction of mosquito larva by a factor of two, say, would only half Zo,whereas doubling the adult death rate would affect an exponential reduction in Zo. Moreover, a sporogonic cycle lasts roughly as long as the aquatic stages of development (see Chapter 1, section 1.4). This observation was later to change the strategy of control measures against mosquitoes (Harrison, 1978, Chapter 23).

2.2.2 Seasonal Variation of prevalence

It has been long established (Boyd, 1949, Peter and Standfast, 1960) that in regions of endemic malaria, prevalence of the disease has a steady seasonal pattern from year to year. As an illustration, Aron and May (1980) examined hospital records of monthly malaria cases, before control by DDT, from two different localities, in Sri Lanka. They found that for regions of high transmission, the records exhibited steady seasonal patterns over the years, quite distinct from localized fluctuations. Such patterns tend to follow those in mosquito population density. The peak of mosquito population density occurs either before (Boyd, 1949, page 636) or during (Christophers, 1949, page 703) the peak of human malaria cases. According to Boyd, the peak of prevalence among mosquitoes is preceded by both the human

and mosquito density peaks.

Aron and May (1980) have considered incorporating this seasonal effect into the basic model of Ross. They introduced a third differential equation in addition to (1) and (2) that describes the dynamics of the mosquito population. In our notation, this third equation is equivalent to

$$\frac{da}{dt} = w(t) - \nu' a \qquad (9)$$

where they define w(t) as the "rate of emergence" of mosquitoes per person. w(t) is made to vary sinusoidally over the year so that the mosquito density, now a function of time, t, also varies sinusoidally but with some lag. Since only the human population is now assumed constant, equation (1) remains unchanged while equation (2) is modified accordingly, and takes the form

$$\frac{du}{dt} = b' f x(a-u) - (\nu' + \frac{da}{dt}/a)u \qquad (10)$$

Graphical representations of the solutions to which the system of equations (1) (9) and (10) tend, for large t, exhibit qualitative agreement with the observations of Boyd mentioned earlier, with regards to the relative timing of the peaks.

2.3 Superinfection

The model of Ross assumes that an infected person is not re-infected. However, it is generally known that, in areas of endemic malaria, human hosts may harbour simultaneously infections by different species (Cohen, 1973; Molineaux and Gramiceia, 1980, pp.134-9). Furthermore, not only can different strains of a single species flourish side by side in the human

host, but the same strain may re-infect the host should the re-infection occur before sufficient body resistance is developed against the strain. These types of infections are examples of the phenomenon of 'superinfection'.

Although Ross (1916) had mentioned the effect of superinfection, he ignored it entirely in his model, thus assuming secondary infections as lost. Macdonald (1950a) noted the omission and proposed a modification to the model in which infections would arrive in the body and run their course independent of each other. Curiously, the model that ultimately emerged is one in which successive infections queue up to express themselves only when the previous infection has run its full course (Macdonald 1957, 1973; Fine 1975; Bailey, 1975 page 318). Dietz has since proposed a model that embraces Macdonald's original idea (Bailey, 1975 pp.318-20). Aron and May (1980) have mentioned constructing some models of superinfection in which re-infections are successively more easily shed rather than being totally independent of each other. These models are said to lie between those of Dietz and Ross. We shall propose another superinfection model, which is a modification of Dietz's model, in which the rate of a re-infection decreases as the number of infections (broods) in the body increases.

To incorporate the models of superinfection we write the basic equation (1) in the general form:

$$\frac{dx}{dt} = h(1-x) - Rx \qquad (11)$$

where h denotes the proportion of the population receiving infectious bites per unit time ($h \equiv b'f'u$); it is also called

the inoculation rate. R is the rate of recovery from all infections in the body. The various models define R as follows:

$$\text{Ross} : R = \gamma \tag{12}$$

$$\text{Macdonald} : R = \begin{cases} \gamma - h & \gamma > h \\ 0 & \gamma \leq h \end{cases} \tag{13}$$

$$\text{Dietz} : R = h/\{ \exp(\frac{h}{\gamma}) - 1\} \tag{14}$$

where γ is now defined as the elimination rate of an infection from a single brood.

Equation (14) has been arrived at by modifying Ross's original model into a hybrid model in the sense that superinfection is treated as a homogeneous stochastic immigration-death process (Bailey, 1975, pp.319-325). As an alternative, if we view superinfection as a queueing process in equilibrium, with both the arrival (infection) and service (recovery)times exponentially distributed, then both equations (13) and (14) are simply special cases in which broods are assumed to arrive and depart according to the queue M/M/S, to use Kendall's notation for queueing processes (Kleinrock, 1975, page 399; Takacs, 1962, page 160). Macdonald's model assumes S, the number of servers, equals 1 (M/M/1), while that of Dietz corresponds to an infinite number of servers, that is $M/M/\infty$.

To obtain a general expression for R, the overall recovery rate, we proceed as follows. Let p_k denote the proportion of the human population with exactly k broods, $k \geq 0$. p_k may also be viewed as the probability of a human host having exactly k broods. In time δt, the proportion of

the population which completely recovers is given by $\gamma p_1 \delta t + o(\delta t)$, since only persons with one brood can move by one step to the susceptible state in the infinitesimal time δt. Now, the proportion of the population with at least one brood is $1 - p_o$, so that the overall recovery rate is given by (Bailey, 1975, page 325).

$$R = \frac{\gamma p_1}{1-p_o} \qquad (15)$$

Thus both models are expressible in the form (15), differing only in their expression for p_o,

For the M/M/1 queue based model, we have $p_o = 1-h/\gamma$ whenever $h < \gamma$, otherwise $p_o = o$. The inequality $h < \gamma$ is the well-known necessary and sufficient condition for ergodicity in the M/M/1 queue (Kleinrock, 1975, p.95). We also have $p_1 = p_o(\frac{h}{\gamma})$. Equation (13) is obtained by substituting for p_o and p_1 in (15). Similarly, we note that, in the case of M/M/∞, we have $p_o = e^{-h/\gamma}$ and p_1 also equals $p_o(\frac{h}{\gamma})$, which if substituted in (15) yields equation (14).

2.3.1 A modified Dietz Model

As stated earlier, we now propose a modification to Dietz's superinfection model. We assume that the occurrence of re-infection becomes more difficult as more broods get established in the body.

The modified model is an example of a queue of type M/M/∞ with discouraged arrivals. One specific and simple case is that of an harmonic discouragement of re-infections with respect to the number of broods present in the body.

Thus, p_k now becomes

$$p_k = p_o \prod_{i=o}^{k-1} \frac{h/(i+1)}{(i+1)\gamma}$$

$$= p_o (\frac{h}{\gamma})^k \frac{1}{k!\ k!} \quad , \qquad (16)$$

where

$$p_o = 1/\{1+ \sum_{k=1}^{\infty} (\frac{h}{\gamma})^k \frac{1}{k!\ k!}\}$$

$$= I_o^{-1}(2(\frac{h}{\gamma})^{\frac{1}{2}}) \quad , \qquad (17)$$

where $I_o(.)$ is the modified Bessel function of the first kind with order zero. So, if we use the new expressions (16) and (17) for p_1(k=1) and P_o respectively in equation (15) we obtain

$$R = h/\{I_o(2(\frac{h}{\gamma})^{\frac{1}{2}}) - 1\} \qquad (18)$$

Our model of superinfection falls somewhere between that of Dietz and Ross, as we shall illustrate subsequently. It is interesting to note that, in this respect the model is similar to those mentioned by Aron and May.

The expressions of Zo corresponding to the various models, when h is replaced by $b'ue^{-\nu'\tau}$, are as follows :

$$\text{Macdonald :} \quad Zo = \begin{cases} \dfrac{ab'^2 f}{\nu'(\gamma e^{\nu'\tau} - b'u)} & \gamma > h \\ \infty & \gamma \leq h \end{cases} \qquad (19)$$

$$\text{Dietz} \quad : \quad Zo = \frac{ab'f}{\nu'u}\{\exp(\frac{b'ue^{-\nu'\tau}}{\gamma}) - 1\} \qquad (20)$$

$$\text{Our Model} : \quad Zo = \frac{ab'f}{\nu'u}\{I_o(2(\frac{b'u}{\gamma})^{\frac{1}{2}} e^{-\frac{\nu'\tau}{2}}) - 1\} \qquad (21)$$

A cursory inspection of equations (19), (20) and (21) indicates that the qualitative conclusion reached by Macdonald about insecticides being more effective than larvicides (as control measures against mosquitoes), can still be drawn from either of expressions (20) and (21). It is worth noting that mosquito density no longer enters linearly in any of the three expressions of Zo. This becomes clearer when u is replaced by ay in these expressions, where, as before, y denotes the proportion of infected mosquitoes. However, because of the double exponential expression of ν' in (20) and the exponential behaviour of $I_o(\cdot)$ for large arguments, (Arfken, 1970, page 511), changes in Zo would tend to be greater in the last two equations than that in (19), given the same change in say a or ν'.

2.3.2 Discriminating Among the Models

The difficulty of discriminating among these models has been pointed out by Aron et al (1980). They suggest one reason for the difficulty to be the observation that under conditions of constant infection rate and constant brood elimination rate, γ, the parameters h and R are related in an identical way for all the models. To illustrate this, they solve equation (11) to obtain the age-prevalence curve x(t), which is

$$x(t) = L(1-e^{-Kt}) \qquad (22)$$

where $L \doteq h/(h+R)$ is the limiting prevalence rate and $K \doteq h+R$, with R defined by either of equations (12), (13) and (14) (and now also (18)), of the various models. Now, irrespective of any of these models, L and K are related by the equation

$$h = LK \qquad (23)$$

One implication of (23) is clearly that, given any two of the quantities h, L and K, a value of γ can be computed which is dependent on the model used. For illustration, some results of estimating h and γ for the models of Ross, Dietz and Macdonald are reproduced (Aron et al., 1980) in Table 1. These estimates are based on the values of L and K obtained by Macdonald (1973, pages 104-106) using equation (22) to fit age-prevalence data from Freetown and Kissy, both in Sierra Leone. We have added to the content of the table (last column) the corresponding estimates of γ as computed by our model.

Clearly the new estimates of γ lie between those from Dietz's and Ross's models. It may also be noted that the new model appears to provide more stable estimates of γ for high values of L than, say, that of Dietz. That all the models at low values of L approach the model of Ross in which superinfection is excluded, is to be expected. They all reflect the negligible occurrence of superinfection at such low levels of disease prevalence.

The general relationship between γ and the other parameters L and h, for the various models may be further illustrated as in Figure 4. The figure is the result of plotting γ/h (the horizontal axis), the relative elimination rate of a brood, against L (the vertical axis), the limiting prevalence rate. Again it is clearly seen that our proposed model lies between the models of Ross (thick unbroken line) and Dietz (thin unbroken line). The figure also shows that all models are bounded from below by the model of Ross and from above by that of Macdonald.

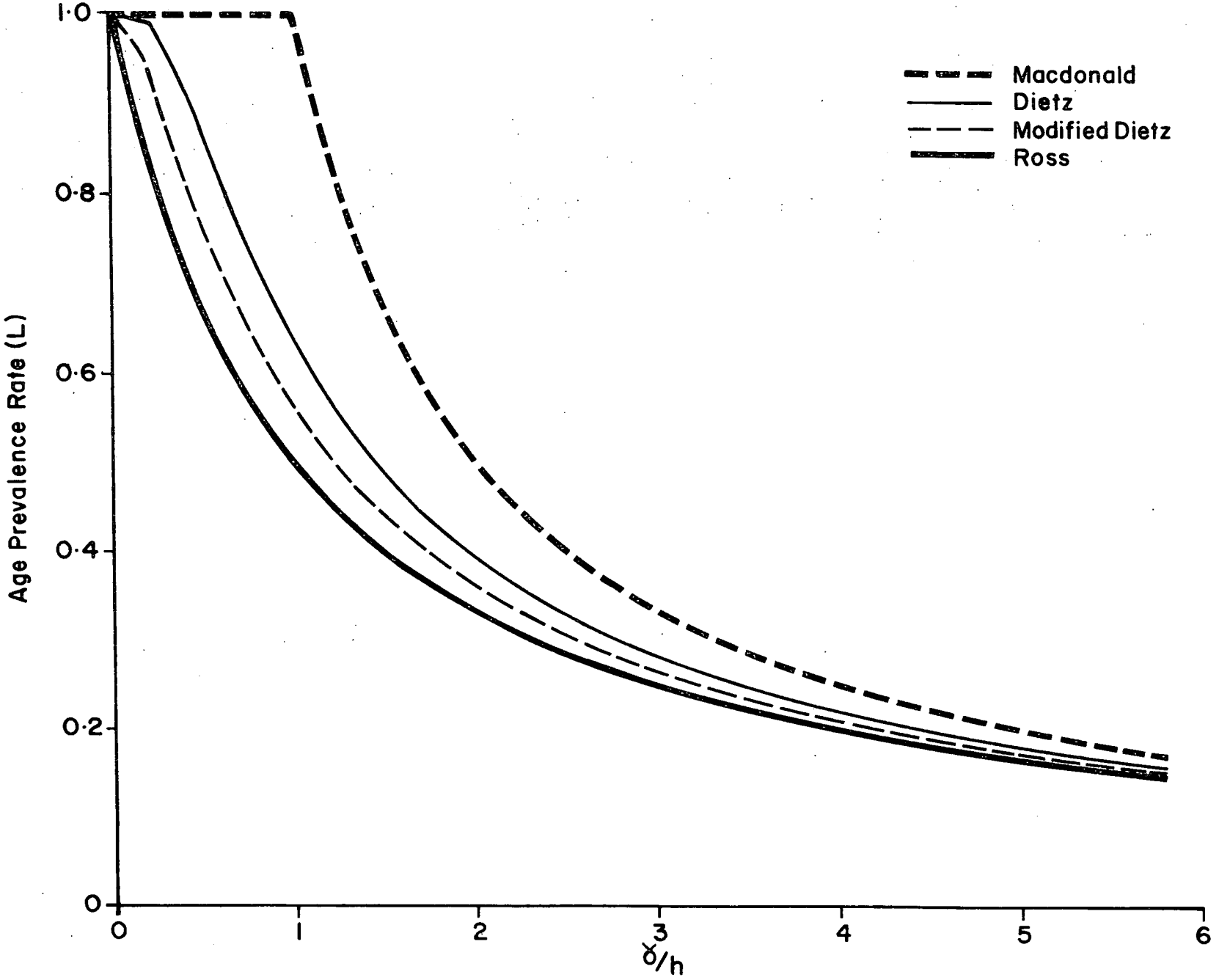
TABLE 1

DATA			Infection rate	Elimination rate $\gamma(yr^{-1})$			
Place & Date	Value of L	Value of $K(yr^{-1})$	$h=KL(yr^{-1})$ for all models	Macdonald	Dietz	Ross	New model
Freetown, 1925	0.41	0.005	0.0021	0.005	0.0039	0.0029	0.0034
Freetown, 1933	0.88	0.005	0.0044	0.005	0.0021	0.0006	0.0013
Freetown, 1947	0.065	0.005	0.0003	0.005	0.0048	0.0047	0.0048
Kissy, 1933	1.0	0.013	0.0130	0.0130*	0	0	0
Kissy, taking L to be 99%	0.99	0.013	0.0129	0.0129	0.0028	0.0001	0.0013

Values of infection rate h and brood elimination rate γ, under the four assumptions about superinfection expressed by equations (10)-(12) and (18).

* γ is indeterminate for L=1.0.

Fig 4. RELATION BETWEEN THE INTRINSIC RECOVERY RATE γ, AND h AND L FOR VARIOUS MODELS OF SUPERINFECTION



However, to be able to distinguish among these models we require some independent methods of estimating γ. With such an independently obtained γ we may then use, for instance, Figure 4 to compare the adequacy of the models. This may be carried out by using the discrepancies between the L values as obtained from the graph with those observed from the population, to set up a simple statistical test of either rejecting or accepting a particular model.

There are a number of difficulties in obtaining such independent estimates of γ. Among these are the problems of detectability and relapses, and ethical reasons preventing measurements under laboratory conditions (Aron et al., 1980). If, in addition, we consider the variation among parasite strains and racial groups, then γ becomes a very elusive parameter to estimate with sufficient accuracy.

It has been suggested that an independent estimate of h, on the other hand, may be obtained either from cohort studies of infants in hyperendemic areas (Pull and Grab, 1974; Molineaux and Gramiccia, 1980,pages 125-126), or from known proportions of persons ever infected, expressed as a function of age t. In either case, R is set to zero, so that the solution of equation (9), confined to the relevant group (for instance, infants) becomes

$$x(t) = 1-e^{-ht} \qquad (24)$$

from which we obtain

$$h = \log_e(1-x(t))/t \qquad (25)$$

According to Aron et al. one way to investigate the effect of superinfection is to carry out comparative studies

between two populations in which the brood elimination rates are assumed to be identical. Such studies may be between closely adjacent communities with different transmission rates as in Segal et al. (1974), or, comparisons of the seasonal variations in recovery rates for a single group of individuals as in Krafsur and Armstrong (1978) and the Garki data (for example, Bekessy et al., 1976; also Chapter 6). As will be seen later (Chapter 6), all the models of superinfection tend to predict longer recovery times during high transmission seasons and shorter recovery times during off seasons. However, these statements are at best suggestive and not a consequence of the models.

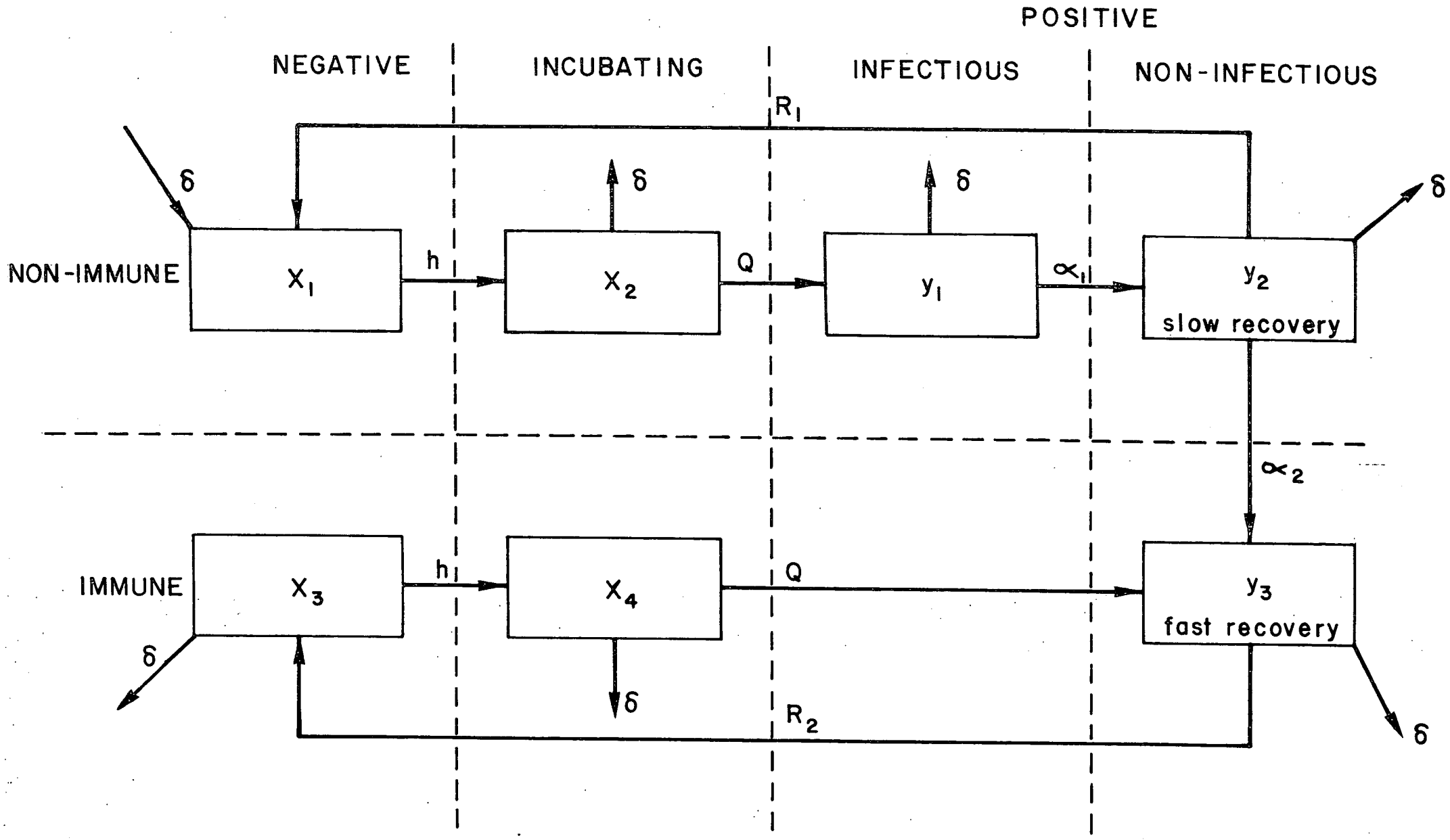
2.3.3 Concluding Remarks

Briefly, the phenomenon of superinfection is important in elucidating the nature of the prevalence of endemic malaria. Its inclusion is therefore imperative in any model vying for realism. However, any attempt to discriminate among the different models of superinfection is beset with the problems of inadequate data. In addition, Macdonald's model is at variance with his original assumption. The model of Dietz which is based on the correct interpretation of Macdonald's original concept, although reasonable, does ignore another important effect of the disease, namely immunity, and possible competition among broods (Molineaux et al., 1980; Singer et al., 1980). The modification we have made to Dietz's model is an effort to take the latter into account. In the next section we review some of the attempts to allow for immunity and, in particular, the integrated malaria model of Dietz.

2.4 The Model of Dietz et al.

The model of Dietz et al. (1974) is considered to be the most practical and realistic yet. It is a deterministic model although some of its assumptions are based on stochastic concepts. It attempts to incorporate the idea of superinfection in the way modelled by Dietz and described in the last section, seasonal variation and the effect of immunity. In discussing the model we shall use the original notations of the authors.

The human population is split into two classes, one immune and the other non-immune. Each class is further divided into several epidemiological categories : negatives or susceptibles, that is persons with no clinically detectable parasites in the body, denoted by x_1 for non-immune and x_2 for the immune; incubating, persons with parasites in the liver only denoted by x_2 and x_4 for the non-immune and immune respectively; the positives, persons with parasites in the blood, denoted by y_1 and y_2 for non-immune, and y_3 for immune classes. The notations x_i, $i = 1,..,4$ and y_i $i = 1, 2, 3$, refer both to the proportions and the categories of the human population at a given time t, and they add up to unity. Here, a unit time is considered as one day. Figure 5 is a diagrammatic representation of the model. The positives consist of infectious, y_1 and non-infections y_2 (non-immune) and y_3 (immune). It is assumed that not all persons with parasites in their blood are capable of making mosquitoes infectious, hence the distinction. The rest of the parameters denote rates of transfers between categories


POSITIVE
NEGATIVE
INCUBATING
INFECTIOUS
NON-INFECTIOUS
R_1
δ
δ
δ
δ
NON-IMMUNE
X_1
h
X_2
Q
y_1
α_1
y_2
slow recovery
α_2
IMMUNE
X_3
h
X_4
Q
y_3
fast recovery
δ
δ
δ
R_2

and are defined as follows (Bailey, 1975, page 324).

δ = birth and death rates of human population for all categories;

h = infection rate (time-dependent) as before;

α_1 = transfer rate from infectious to non-infectious category in the non-immune class;

α_2 = transfer rate from the non-immune (non-infectious) class to the immune (non-infectious) class;

Q = transfer rate from the incubating category of either class to either the infectious (non-immune class) or the non-infectious (immune class) categories; given by $(1-\delta)^N h(t-N)$;

R_1, R_2 = recovery rates of non-immune and immune classes respectively given by $h/\{ \exp(\frac{h}{\gamma_i}-)-1 \}$, $i = 1,2$

Here, N is the incubation period of infection in man and γ_i is the elimination rate of any one brood. From the expression of Q, it is assumed that immunity has no influence on incubation.

The model is formulated as a set of non-linear difference equations amenable to iterative solution by computer. Thus, let $\Delta x_j \equiv \Delta x_j(t) = x_j(t+1)-x_j(t)$, $j = 1,\ldots,4$, and $\Delta y_i \equiv \Delta y_i(t) = y_i(t+1)-y_i(t)$, $i = 1,2,3$, then the equations are given as follows, with the time argument

suppressed except where a time-lag is involved.

$$
\begin{aligned}
&\text{(i)} && \Delta x_1 = R_1 y_2 - (h+\delta)x_1 + \delta ; \\
&\text{(ii)} && \Delta x_2 = h x_1 - Q x_1(t-N) - \delta x_2 ; \\
&\text{(iii)} && \Delta x_3 = R_2 y_3 - (h+\delta)x_3 ; \\
&\text{(iv)} && \Delta x_4 = h x_3 - Q x_3(t-N) - \delta x_4 ; \\
&\text{(v)} && \Delta y_1 = Q x_1(t-N) - (\alpha_1+\delta)y_1 ; \\
&\text{(vi)} && \Delta y_2 = \alpha_1 y_1 - (\alpha_2 + R_1 + \delta)y_2 ; \\
&\text{(vii)} && \Delta y_3 = \alpha_2 y_2 + Q x_3(t-N) - (R_2+\delta)y
\end{aligned}
\qquad (26)
$$

The expression for Q is derived as follows. Transfers from x_2 to y_1 at time t consist of persons infected at time t-N with infection rate h(t-N). If the death rate is δ, then the proportion of survivors after time N is approximately $(1-\delta)^N$. Therefore, the proportion transferred to y_1 at time t is $(1-\delta)^N h(t-N)x_1(t-N)$, which, when equated to the first term of the right side of (26 (v)) gives the required relation. The expressions for R_i, i = i,2, are Dietz's model of superinfection (equation (14)), with γ_1 and γ_2 as the elimination rates of single broods for y_2 (the non-infectious positive non-immunes) and y_3 (the non-infectious positive immunes) respectively. It is assumed that individuals in y_3 recover faster than those in y_2 so that $\gamma_1 < \gamma_2$

As the infection rate, h(t), is made to depend on the vector population it will change with the season. An explicit expression for h(t) is given in terms of a quantity called vectorial capacity, denoted in the literature by C(t) (Garrett-Jones, 1964; Garrett-Jones and Shidrawi, 1969). C(t) is defined as the total number of bites distributed by those mosquitoes surviving beyond their sporogonic cycle of

duration τ (days), having bitten a single individual on day τ. Using some of the notations defined earlier (section 2), C(t) is given as

$$C(t) = \frac{ab'^2 e^{-\nu'\tau}}{\nu'} \qquad (27)$$

That is, C(t) is the product of the average number of bites per person per day (ab'), the proportion of mosquitoes surviving for τ days ($e^{-\nu'\tau}$), the average life expectancy of survivors ($^1/\nu'$) and the average number of bites of a surviving mosquito per person per day (b'). Note that the mosquito density, a, is time-dependent.

An approximate expression for h is given by

$$h(t) = g\{1-\exp(-C(t-N)y_1(t-N))\} \qquad (28)$$

where g is interpreted to mean the probability that a bite actually infects a susceptible person. The expression in the curly bracket is derived as follows : Infectious persons on day (t-N), ($y_1(t-N)$), make on day t and thereafter (through the mosquito population) an average number, $C(t-N)y_1(t-N)$, of potentially infectious contacts with each member of the human population. By assuming that all these bites occur on day t itself and that their number has a Poisson distribution, the probability that any given susceptible person receives at least one infectious bite is given by $1-\exp\{-C(t-N)Y(t-N)\}$.

Dietz et al. also derived an expression for what they call the critical vectorial capacity, denoted by C^*, below which the disease cannot maintain itself at an endemic level. This is based on setting the reproduction rate Zo to unity. The expression is obtained by finding, in terms of C, the

average number of secondary cases that result from a single primary infection. Thus the mean period during which a case is infectious is $(\alpha_1+\delta)^{-1}$. During this time, for small vectorial capacities the person makes approximately gC successful contacts per unit time. Hence,

$$Z_o \equiv \frac{gC}{(\alpha_1+\delta)} = 1$$

or

$$C^* = (\alpha_1+\delta)/g \qquad (29)$$

As mentioned already, the development of the model by Dietz et al. was closely linked with the Garki project. It has been applied to the baseline data of the project and has provided what the authors consider a good fit to the observed variation in malaria prevalence by age and season of the area. It has also been used more recently to predict the epidemiological pattern of another endemic malarial community in Kenya (Molineaux et al., 1978). It has yet to be applied to epidemiological circumstances far removed from those prevailing in tropical Africa.

2.4.1 Some Extensions

As it stands the model assumes that immunity once gained, is not lost. However, this is far from the case. Mention has been made already of the fact that immunity wears off gradually when transmission in an area is significantly reduced by use of some form of control (Chapter 1, Section 1.5). Moreover, immunity is also reduced when a person remains uninfected for some time (McGregor et al.,1966).

A model of loss of immunity has been suggested by Aron et al. (1980), in which they allow immunity loss to decrease with increasing value of the infection rate. The

derivation has some parallel with Dietz's superinfection model and it is of the form:

$$\lambda = h/ \{ \exp(hT) - 1 \} , \qquad (30)$$

where λ is the rate of moving out of the immune to the non-immune state, and T denotes the mean time immunity lasts in the absence of re-infection. Equation (30) is incorporated into a simple SIRS (Susceptible-Infected-Recovered/Immune-Susceptible) model, represented by a set of linear differential equations. They make a study of the solutions to these equations under various assumptions about infection rates and immunity levels. Their result indicates that gradual reduction of transmission when initial rates of infection is high, may diminish the level of naturally acquired immunity in adults and actually raise their prevalence level before allowing it to fall as h continues to fall. Real age-prevalence curves studied by Boyd (1949) exhibit trends similar to the theoretical results of Aron et al. Clearly, this situation may have serious implications for control programmes because such a rise in prevalence curve may be indistinguishable from that due to actual lack of success of the application of a control measure.

Another shortcoming of the model by Dietz et al. is the assumption that immune persons, when infected, never become infectious. We have already mentioned that acquired immunity plays almost no part in the removal of gametocytes (see Section 1.5). It does, however, affect indirectly the density of gametocytes in the blood since it removes some of the merozoites which would otherwise give rise to

fresh gametocytes. Even so, according to Yekutiel (1960) mosquitoes can get infected by gametocytes occurring in such small numbers as to escape detection easily. One way we would include the contribution of all infectious persons is as follows.

Let y_4 represent both the category and proportion of persons who are infectious and immune, and let $\gamma_3 (> \gamma_1)$ denote their transfer rate into the non-infectious but immune state, y_3. If q is the ratio of gametocyte density of infectious immunes to those of infectious non-immunes, or some appropriately chosen parameter, then the infection rate, h, now becomes

$$h(t) = g\{ 1-\exp(-C\ (t-N)\ (y_1(t-N)+qy_4(t-N))) \} \qquad (31)$$

The modification adds one non-linear difference equation to (26) (Figure 6)

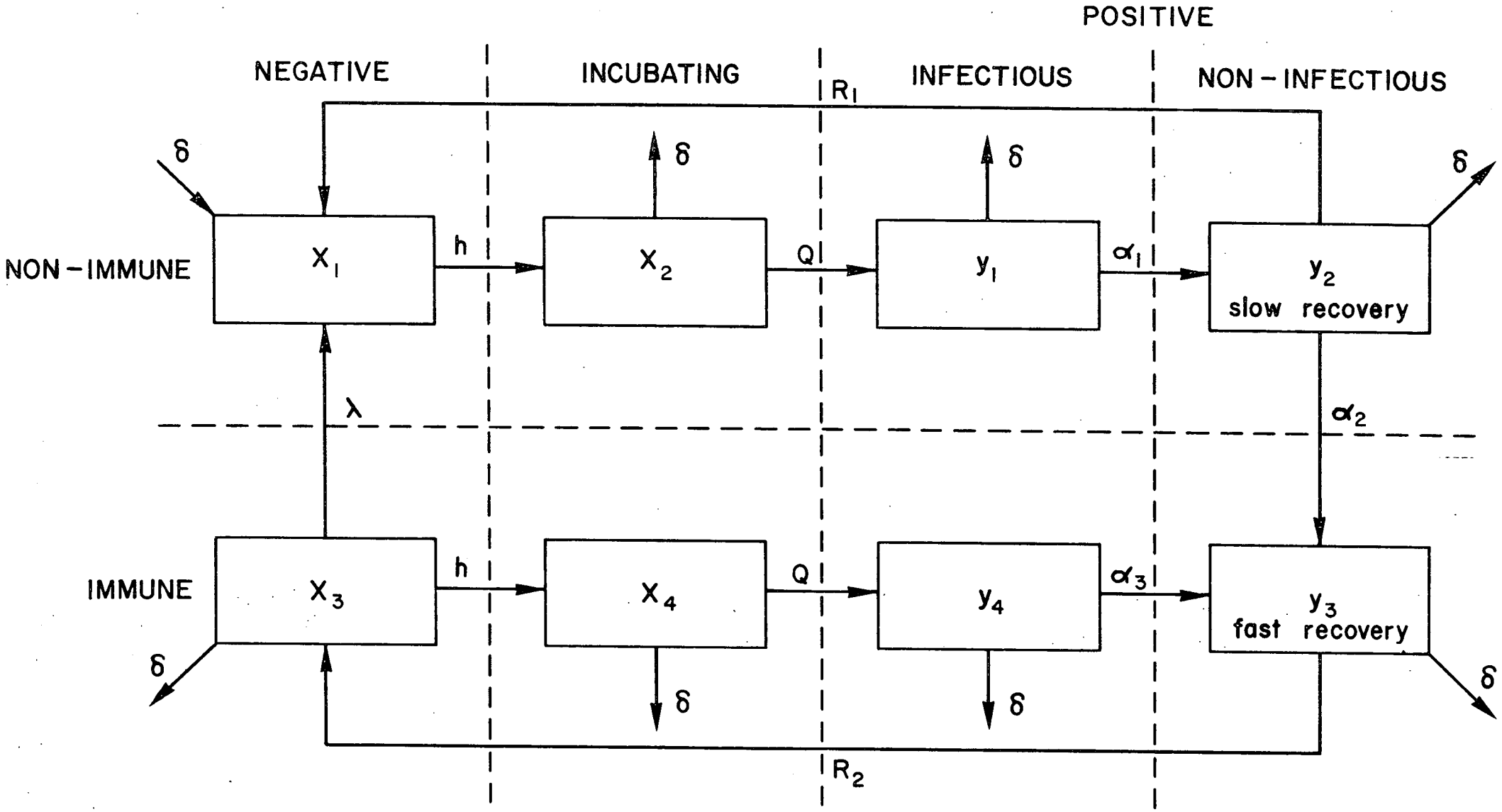
$$\Delta y_4 = Qx_2(t-N) - (\gamma_3+\delta)y_4 \qquad (32)$$

In addition equation (26 (vii)) now becomes

$$\Delta y_3 = \gamma_2 y_2 + \gamma_3 y_4 - (R_{\geq 2} +\delta)y_3 \qquad (26\ (vii)')$$

In general, $q < 1$, a consequence of the reduction of merozoite density due to immunity. Clearly, expression (31) provides an estimate of h which is consistently higher than that of Dietz et al. Thus, h reaches saturation level at slightly lower vectorial capacity than in (28).

Fig 6 STATES AND TRANSITIONS OF MODIFIED DIETZ MODEL



In endemic areas and more so during high transmission period, y_4 is bound to be high and its effect may not be negligible as assumed by Dietz et al. One implication is that a lot more effort may be required at the initial stages of a control programme before a noticeable change is realized in h.

The parameter q in equation (31) is a simple way of incorporating the magnitude of the parasite burden in the human population. It has been argued that since the population of malaria parasites can exist at various densities within infected persons, then the epidemiological manifestations of the disease depend on these quantitative aspects of infection rather than on mere presence or absence of parasites in the body. Besides, it is a common occurrence that, because of the cyclic nature of the development of the parasites in the blood, some negative results of examination of blood slides may be false (Boyd 1949; Miller, 1958). It would be interesting to see how significant an improvement these modifications would bring about in the model.

2.5 The Continum Model

The model of Elderkin et al. (1977), which is based on the suggestions of Dietz in cooperation with other malaria epidemiologists at WHO, is constructed to take into account the aspects of malaria described in the last paragraph of the preceding section. The model attempts to describe the dynamics of malaria in a host rather than its transmission in a community as in previous models. In the model, the immune response and the magnitudes of the burden of the various forms of the parasite in the human host are made to be both age- and time-dependent. But, it is the time independent (age-dependent) version that the authors consider in some detail.

The set of equations that define the model is as follows:

$$\text{(i)} \quad \frac{dy}{dt} = V \int_o^\infty \delta e^{-\delta v} g(v)dv - \rho y \; ;$$

$$\text{(ii)} \quad \frac{dg}{dt} = \alpha y e^{-\rho T} - \nu g \; ;$$

$$\text{(iii)} \quad \frac{d\rho}{dt} = \mu y - \gamma(\rho - \rho_o) \; ; \qquad (33)$$

where, y is the population density of asexual blood stages (merozoites) of the parasites in the host, g is the density of gametocytes in the blood, and ρ is the death rate of the asexual forms or level of resistance of the host. The three variables are all dependent on the age of the host, denoted here by t. The parameters in the model have the following meaning.

V : the "biting rate" of the mosquitoes and is interpreted similarly to C, the vectorial capacity in the model of Dietz et al.

δ : human mortality rate ;

α : the rate of reproduction of gametocytes ; (cf. α_1 and α_2 in section 2.5) ;

ν : the death rate of gametocytes ;

μ : rate of increase in level of immunity ;

γ : rate of loss of immunity.

The equations are interpreted as follows. The change in y depends on g, the density of gametocytes transferred to mosquitoes given by $V\int_0^\infty \delta e^{-\delta v} g(v)dv$. This is the product of the 'biting rate' V, of the mosquitoes and the density of gametocytes in all persons surviving to an age v, $\int_0^\infty \delta e^{-\delta v} g(v)dv$. The loss term, ρg, is the death of merozoites due to the immune response of the host. As for equation (33 (ii)), increases in g are proportional to the merozoites conditional on their surviving for a time T, where their approximate probability of survival is given by $e^{-\rho T}$. Gametocytes are also reduced at a constant rate ν, by νg. In the last equation (33 (iii)), ρ increases proportionally with y at a constant rate μ. In the absence of merozoites ($y=o$), ρ is made to decrease at a constant rate γ to a value ρ_o, the immunity level of a new-born child, hence the term $\gamma(\rho - \rho_o)$.

Conditions for the existence of non-trivial solutions to these equations is summarized in their theorem 1.1. That is, given the initial conditions:

$$\begin{aligned} &(i) \quad y(o) = y^o \geq o \ ; \\ &(ii) \quad g(o) = g^o \geq o \ ; \\ &(iii) \quad \rho(o) = \rho^o \geq o \end{aligned} \tag{34}$$

then the set of equations (33) has at least one solution which may be either trivial ($y = g = \rho = o$) or non-trivial depending on the following conditions :

(a) $y^o = g^o = o$; there is only a trivial solution

(b) $V > \alpha^{-1}(\nu+\delta)\ (\rho_o+\delta)e^{\max\ (\rho_o, \rho^o)T}$;

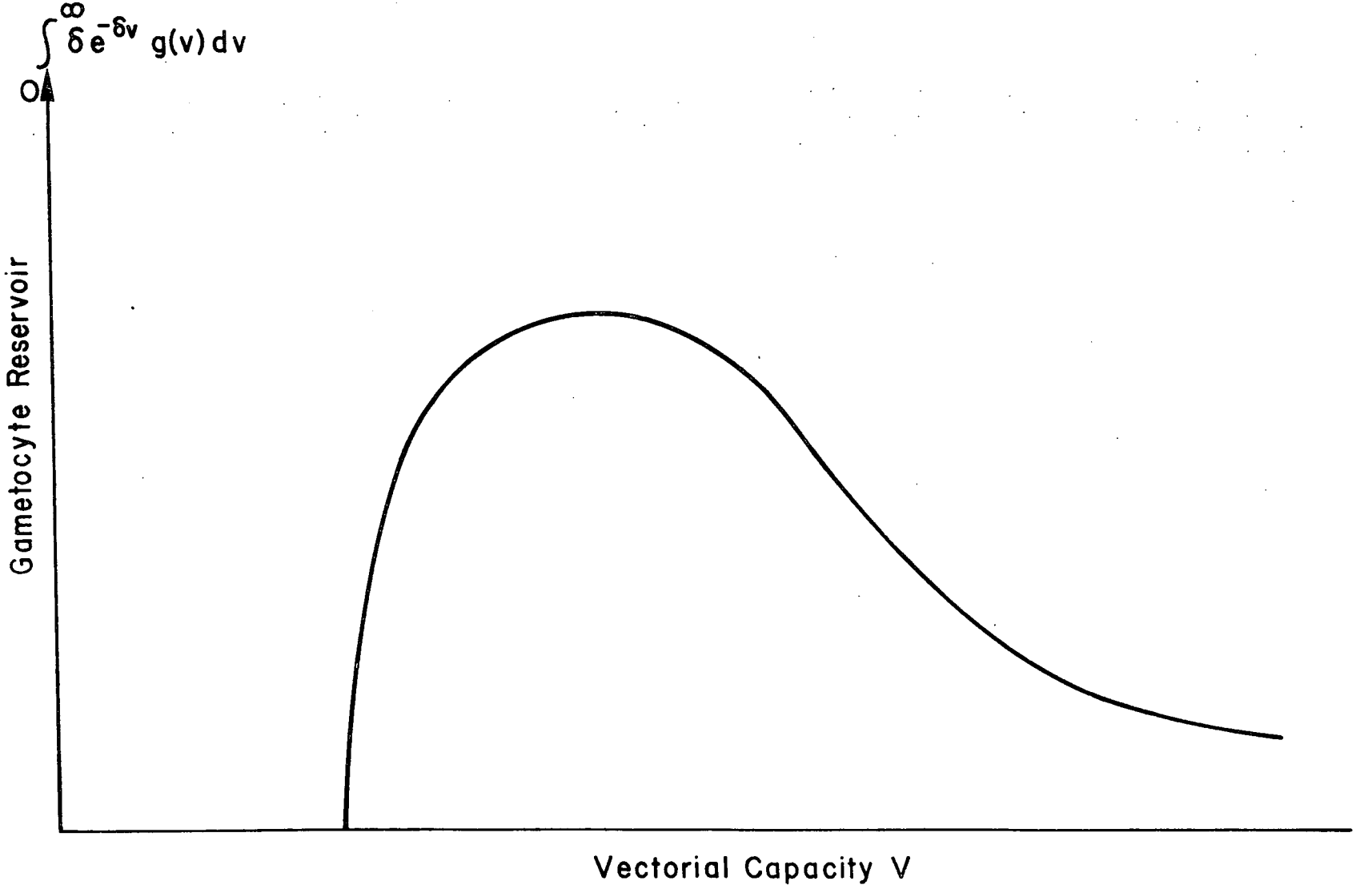
there is a non-trivial solution, otherwise there is a trivial solution if

$V \leq \alpha^{-1}(\nu+\delta)(\rho_o+\delta)e^{\min(\rho_o, \rho^o)T}$

An important feature of the model is that it allows for resistance to be re-inforced only in the presence of asexual blood forms of the parasites. It also allows for it to erode in the absence of the asexual stages of the parasite. These assumptions are in agreement with empirical evidence outlined earlier (Chapter 1, subsection 1.5.1) with regard to the process of the acquisition and loss of immunity.

The main epidemiological conclusion of the model is that gametocyte reservoir increases with increasing transmission at low vectorial level (V), but that, as immunity builds up gradually with increasing transmission the reservoir eventually subsides (Aron and May, 1980). This property is exhibited by Figure 7, reproduced from Elderkin et al. The interpretation of this result could be that a moderate reduction in transmission from a high transmission level is likely to increase the prevalence of the disease among adults, that is, things will get worse before getting better.

Fig. 7 GAMETOCYTE RESERVOIR AS A FUNCTION OF THE VECTORIAL CAPACITY



Source : Elderkin et al, 1977

CHAPTER THREE

ON SEMI-MARKOV PROCESSES

3.1 Introduction

We proceed to sketch some of the mathematical tools and derivations to be used in subsequent discussions. Since the basic model to be considered in the sequel is semi-Markovian, the discussion here will focus on the theory of semi-Markov processes, and in particular, on those features of the theory deemed pertinent later on.

The ideas of semi-Markov processes were independently conceived by Lévy (1954) and Smith (1955). About the same time, Takács (1954) introduced essentially the same type of stochastic processes and applied them to some problems in counter theory. Since then, various aspects of the process have been investigated, under the general name of Markov renewal processes, by Pyke (1961a,b), Taga (1963), Pyke and Schaufele (1964) and Moore and Pyke (1968). Recent expositions of the theory are by Cinlar (1969, 1975) and Hunter (1969). Weiss and Zelen (1965) discuss parameter estimation problems of models based on semi-Markov theory in the context of application to clinical trials. Other workers on parameter estimations include Lagakos et al. (1978) (on right-censored data) and Thompson (1981). An example of application in the social sciences is that of Ginsberg (1971a), who uses it to model social mobilities in the United States. Ginsberg (1971 b, c) has also made a critical appraisal of the application of semi-Markov theory in modelling migration in general.

In section 3.2 we give basic definitions and introduce some notations. We derive the transition probabilities of the process in section 3.3 followed by their limit values in section 3.4. Section 3.6 contains the only new results in this chapter, the joint limiting distributions. These are derived using the formula of forward recurrence time distributions outlined in section 3.5. A simple illustrative example is given in section 3.7. And we conclude with a few remarks (section 3.8) about the appropriateness of semi-Markov models in general.

3.2 Definitions and Notations

We consider a stochastic process which moves from one to another of a number of states 1, 2, ..., N, in which successive states visited, denoted by the set $\{ X_o, X_1, \ldots \}$, form a discrete time Markov chain with transition probabilities $\{ p_{ij} \}$, $i,j = 1, 2, \ldots, N$. Here, X_o denotes the initial state of the process and X_n, $n = 1, 2, \ldots$, the state immediately following the nth transition so that $X_n \varepsilon \{1, 2, \ldots, N\}$. Further, let $\{T_n, n = 1, 2, \ldots\}$ be positive random variables such that

$$\text{Prob} \{ T_n < t \mid X_o, X_1, \ldots, X_{n-1} = i, X_n = j, T_1, T_2, \ldots, T_{n-1} \} = Hij(t)$$

$$\text{for all } n = 1, 2, \ldots \qquad (1)$$

$Hij(\cdot)$ is assumed to be an honest probability distribution.

Definition 1. Let $S_n = \sum_{k=1}^{n} T_k$, $S_o = 0$, define $Z_t = X_n$ for $S_n < t < S_{n+1}$ and suppose $S_n \to \infty$ a.s. Then the process $\{Z_t, t > o\}$ is called a semi-Markov process (Cox and Miller, 1977, page 352; Pyke, 1961a).

The number of states of the process, N, may be either finite or infinite. In our case, N is assumed finite.

The process is also characterized by an N x 1 vector of initial probabilities, $\underline{A} = (a_1, a_2, \ldots, a_N)$, defined by $a_i = \text{Prob}\{ X_o = i \}$, $i = 1, 2, \ldots, N$, and satisfying (i) $a_i \geq o$, $i = 1, 2, \ldots, N$, (ii) $\sum_{j=1}^{N} a_i = 1$.

We define $F_{ij}(t) = p_{ij}H_{ij}(t)$, and $W_i(t) = \sum_{j=1}^{N} F_{ij}(t)$. Since $H_{ij}(\cdot)$ is assumed honest, we have

$$\lim_{t\to\infty} W_i(t) = \sum_{j=1}^{N} \lim_{t\to\infty} F_{ij}(t) = 1 \qquad i = 1,2,\ldots,N \tag{2}$$

Consequently we may define $p_{ij} = \lim_{t\to\infty} F_{ij}(t)$.

Let $\phi_{ij}(\cdot)$ be defined for all i,j and $t \geq o$, by

$$\phi_{ij}(t) = \text{Prob}\{ Z_t = j \mid Z_o = i \} \tag{3}$$

Thus, $\phi_{ij}(t)$ is the probability that the process, initially in state i, is in state j at time t.

Unless stated otherwise, matrices of any doubly subscript quantities will be denoted by their corresponding letters underlined. For example, $\underline{F}(\cdot) = (F_{ij}(.))$ The corresponding Laplace transforms will be distinguished by superscript (*). Thus for $\theta \geq o$, $f^*(\theta) = \int_o^\infty e^{-\theta t} f(t)dt$. In the case of Laplace - Stietjes transforms we shall use (**). That is, $f^{**}(\theta) = \int_o^\infty e^{-\theta t} df(t)$. In addition, wherever convenient, functions may be written with their arguments suppressed, provided this causes no ambiguity.

The following additional notations will also be used in the sequel. They are defined for all i,j and $t \geq o$, and are assumed finite :

$$\text{(i)} \quad \mu_{ij} = \int_o^\infty t dH_{ij}(t)$$

$$\text{(ii)} \quad \mu_i = \int_o^\infty t dW_i(t)$$

$$\text{(iii)} \quad \sigma_i^2 = \int_o^\infty (t-\mu_i)^2 \, dW_i(t) \qquad (4)$$

It is clear from the definition of $F_{ij}(.)$ that

$$\mu_i = \sum_{j=1}^{N} p_{ij} \mu_{ij} \qquad (5)$$

3.3 The Transition Probabilities

The functions $\phi_{ij}(\cdot)$ may be expressed recursively. By the argument of total probability, they are expressed as follows (Çinlar, 1969; Pyke, 1961b; Feller, 1957).

$$\phi_{ij}(t) = \delta_{ij}(1-W_i(t)) + \sum_{k=1}^{N} \int_o^t d F_{ik}(\tau) \, \phi_{kj}(t-\tau) \qquad (6)$$

$$i,j = 1,2,..,N, \quad t \geq \tau \geq o$$

where δ_{ij} is the usual kronecker delta function. Equations (6) are derived as follows. The first term on the right arises from the case i = j and the situation that the process never left state i during the entire time t. The second term is the case in which the process left state i for some intermediate state k, and eventually reached state j at time t.

If we denote the diagonal matrix $(\delta_{ij} W_i(\cdot))$ by $\underline{W}(t)$,

then equations (6) may be written compactly using matrix notations. Thus,

$$\underline{\Phi}(t) = (\underline{I}-\underline{W}(t)) + (\underline{F}*\underline{\Phi})(t) \qquad (7)$$

where $(\underline{F}*\underline{\Phi})(\cdot)$ is the matrix convolution of $\underline{F}(\cdot)$ with $\underline{\Phi}(\cdot)$. To solve for $\underline{\Phi}(\cdot)$, we use the Laplace transform of (7). For $\theta > 0$

$$\underline{\Phi}^*(\theta) = \frac{1}{\theta}\underline{I} - \underline{W}^*(\theta) + \theta\underline{F}^*(\theta).\ \underline{\Phi}^*(\theta). \qquad (8)$$

Solving for $\underline{\Phi}^*(\theta)$, equation (8) yields

$$\Phi^*(\theta) = \frac{1}{\theta}\,(\underline{I}-\theta\underline{F}^*(\theta))^{-1}(\underline{I}-\theta\underline{W}^*(\theta)) \qquad (9)$$

The inverse in the right is well defined since $F_{ij}(0) = 0$ and N is finite (Pyke 1961b).

The whole process depends only on the p_{ij}'s, $H_{ij}(\cdot)$'s and the initial probability distribution $\underline{A}$. In particular, given $F_{ij}(\cdot)$'s (or p_{ij}'s and $H_{ij}(\cdot)$'s),in theory, equation (9) can be used to obtain $\Phi_{ij}(\cdot)$'s. This is significant in that it provides the means to express $\Phi_{ij}(\cdot)$'s in terms of the parameters of the underlying process, which may be useful for statistical analysis.

3.4 Limiting Behaviour of $\Phi_{ij}(\cdot)$

One other useful property of $\Phi_{ij}(\cdot)$ is its limiting form as t tends to infinity. In matrix notation, let $\underline{\Phi}$ be defined by

$$\underline{\Phi} = \lim_{t\to\infty} \underline{\Phi}(t) \qquad (10)$$

where the limit operation is carried on each element of $\underline{\Phi}(\cdot)$. In terms of Laplace transforms these limits are given

by (Cox and Miller, 1977, page 355)

$$\underline{\Phi} = \underset{\theta \to o}{\text{limit}} \{\theta \underline{\Phi}^*(\theta)\} \tag{11}$$

A rearrangement of equation (9) gives

$$\underline{\Phi} = \lim_{\theta \to o} \{\theta \underline{\Phi}^*(\theta)\}$$

$$= \lim_{\theta \to o} \theta(\underline{I} - \theta \underline{F}^*(\theta))^{-1} x \lim_{\theta \to o} \frac{1}{\theta} (\underline{I} - \theta \underline{W}^*(\theta)) \tag{12}$$

provided these separate limits exist.

Since $\underline{W}^{**}(\theta) = \theta \underline{W}^*(\theta) - \underline{W}(O) = \theta \underline{W}^*(\theta)$, the second term on the right becomes $\lim_{\theta \to o} \frac{1}{\theta} (\underline{I} - \underline{W}^{**}(\theta))$.

Since this expression becomes indeterminate for $\theta = o$ ($\underline{W}^{**}(o) = \underline{I}$), we need to use L,Hospital's rule to obtain the limit. Thus

$$\lim_{\theta \to o} \frac{1}{\theta} (\underline{I} - \underline{W}^{**}(\theta)) = \frac{d}{d\theta} \underline{W}^{**}(\theta) \Big|_{\theta = o}$$

$$= \int_o^\infty t \, d\underline{W}(t)$$

$$= \underline{M} \tag{13}$$

where $\underline{M} = (\delta_{ij} \mu_i)$, the diagonal matrix whose positive elements are the unconditional mean holding times of the states of the process.

As for the limit of the inverse term, we put (Howard, 1963) (noting that $\underline{F}^{**}(\theta) = \theta \underline{F}^*(\theta) - \underline{F}(o)$ and $\underline{F}(O) = o$),

$$\underline{U}^*(\theta) = \theta(\underline{I} - \underline{F}^{**}(\theta))^{-1} \tag{14}$$

or

$$\underline{U}^*(\theta) - \underline{U}^*(\theta)\underline{F}^{**}(\theta) = \theta \underline{I} \tag{15}$$

Now, the matrix $\underline{U}^*(\theta)$ exists by virtue of the existence of the inverse term in (9). Taking limits of (15) as $\theta \to o$ yields

$$\underline{U}^*(o) = \underline{U}^*(o)\,\underline{P} \tag{16}$$

where the transition matrix, $\underline{P}$, is substituted for $\underline{F}^{**}(O)$.

If we assume that the embedded Markov chain of the process is ergodic, then, its limiting state probabilities are unique and independent of the initial state (Chung, 1968; Cox and Miller, 1977, page 101). This result is easily extended to the case in which the process is irreducible positive-recurrent and periodic (Feller, 1957, page 321).

Let $\underline{\Pi}$ denote the row vector of the limiting probabilities of the embedded Markov chain. This vector must satisfy the equations

$$\begin{aligned} &\text{(i)} \quad \underline{\Pi} = \underline{\Pi}\underline{P} \\ &\text{(ii)} \quad \sum_{j=1}^{N} \Pi_j = 1 \end{aligned} \tag{17}$$

where $\Pi_j \geq o$, $j = 1, 2, \ldots, N$, are the elements of the vector $\underline{\Pi}$. Because of the uniqueness of the solution of (17) the rows of $\underline{U}^*(o)$ must each be proportional to the row vector $\underline{\Pi}$.

To obtain the proportionality term the results from equations (13) and (16) are substituted into (12) to get

$$\underline{\Phi} = \underline{U}^*(o)\underline{M} \tag{18}$$

The elements of $\underline{\Phi}$ must be of the form

$$\Phi_{ij} = U^*_{ij}(o)\,\mu_j = k_i \Pi_j \mu_j \tag{19}$$

where $u^*_{ij}(o)$ is the (i,j)th element of $\underline{U}^*(o)$, and k_i is a constant of proportionality between the ith row of $\underline{U}^*(o)$ and $\underline{\Pi}$.

Since $\sum_{j=1}^{N} \phi_{ij} = 1$, then

$$k_i = 1/\sum_{j=1}^{N} \Pi_j \mu_j \tag{20}$$

which are identical for all states.

Hence

$$\phi_{ij} = \frac{\Pi_j \mu_j}{\sum_{\ell=1}^{N} \Pi_\ell \mu_\ell} \tag{21}$$

The above results are summarized in the following

Theorem 1. Let $\{ Z_t ; t \geq o \}$ be an irreducible persistent and aperiodic semi-Markov process. If $\mu_j < \infty$, $j \geq 1,2,\ldots,N$, then

$$\Phi_{ij}(t) \to \frac{\Pi_j \mu_j}{\sum_{\ell=1}^{N} \Pi_\ell \mu_\ell} \quad \text{as } t \to \infty \tag{22}$$

where the row vector, $\underline{\Pi}$, with elements Π_j, is the unique solution of $\underline{\Pi} = \underline{\Pi}\,\underline{P}$, and $\underline{P} = \lim_{t\to\infty} \underline{F}(t)$.

As expected, ϕ_{ij} is independent of the starting point i and depends on the waiting time distribution only through the unconditional mean waiting times μ_ℓ, $\ell = 1,2,\ldots,N$.

3.5 Forward Recurrence Times

We need to define some concepts on Markov renewal process (MRP) and introduce some additional notations. An MRP is the process corresponding to a semi-Markov process which records, at each time t, the number of visits to each of the states of the process up to that time.

More formally an MRP is defined as follows. Define $m(t) = \sup \{ n \geq o \mid S_n \leq t \}$ and, $m_j(t)$, the number of times $X_k = j$ for $1 \leq k \leq m(t)$.

Definition 2 The stochastic process defined by the vector-valued random variable $\underline{M}(t) = \{ M_1(t), \ldots, M_N(t) \}$ is called a Markov renewal process. It reduces to an ordinary renewal process when $N = 1$.

Let $E_{ij}(t)$, $i,j = 1,2,\ldots,N$, denote the Markov renewal function; that is, $E_{ij}(t) = E \{ M_j(t) \mid Z_o = i \}$. The function is usually expressed in terms of $F_{ij}(\cdot)$ (Çinlar, 1969; Pyke, 1961b). Thus

$$E_{ij}(t) = \sum_{k=1}^{\infty} F_{ij}^{(k)}(t) \qquad i,j=1,2,\ldots,N \tag{23}$$

where $F_{ij}^{(k)}(.)$ is the k-fold convolution of $F_{ij}(\cdot)$ with itself.

The following notations are essentially those used by Çinlar (1969). Let $P_j(.)$ denote the probability measure conditioned on $X_o = j$. That is $P_j(.) = \text{Prob} \{ \cdot \mid X_o = j \}$. In addition let

$$n(t) = \sup \{ n, S_n \leq t \}$$

$$V_t^+ = S_{n(t)+1} - t \quad ; \text{ the time until next transition}$$

$$Z_t^+ = X_{n(t)+1} \quad ; \qquad \text{the state to be visited next}$$

The processes $\{X_{n(t)}, V_t^+\}$ and $\{X_{n(t)}, V_t^-\}$ are both Markov whose state spaces are the Cartesian product $\{1,2,\ldots,N\} \times (o, \infty)$. The following results about these processes are largely from Cinlar (1969) and are stated without proofs.

Let $X_o = j$ and consider the event $\{ Z_t = k, V_t^+ \leq y \}$, $(Z_t = X_{n(t)})$. Conditioning on S_1, and invoking the theorem of total probability, we obtain the renewal equations

$$P_j \{ Z_t = k, V_t^+ \leq y \} = \delta_{jk}(W_j(t+y) - W_j(t)) + \sum_{\ell=1}^{N} \int_o^t dF_j(u)$$

$$P_\ell \{ Z_{t-u} = k, V_{t-u}^+ \leq y \} \qquad j,k = 1,2,\ldots,N \qquad (24)$$

Since the first term in the right is bounded, then the solution to (24) is given by (Karlin and Taylor, 1975, pp.184-5)

$$P_j \{ Z_t = k, V_t^+ \leq y \} = \int_o^t dE_{jk}(u) \; (W_k(t+y-u) - W_k(t-u)) \; ; \qquad (25)$$

$$j,k = 1,2\ldots,N$$

If we denote $\eta_i = \dfrac{\Pi_i \mu_i}{\sum_{\ell=1}^{N} \Pi_\ell \mu_\ell}$ then by

the basic renewal theorem, as $t \to \infty$

$$P_j \{ Z_t = k, V_t^+ \leq y \} \to \frac{\eta_k}{\mu_k} \int_o^y (1-W_k(u)) \; du \; ; \qquad (26)$$

$$j,k = 1,2,\ldots,N$$

which is independent of the initial state j.

The following result is immediate from (25)

$$P_i \{ Z_t = j,\ Z_t^+ = k,\ V_t^+ \leq y \} = \int_0^t dE_{ij}(u)\ (F_{jk}(t+y-u) - F_{jk}(t-u))$$

$$i,j,k = 1,2,\ldots,N \ ; \qquad (27)$$

and hence from (26)

$$P_i \{ Z_t = j,\ Z_t^+ = k,\ V_t^+ \leq y \} \to \frac{\eta_j}{\mu_j} \int_0^y (p_{jk} - F_{jk}(u))\ du \qquad (28)$$

$$i,j,k = 1,2,\ldots,N$$

which is again independent of the initial state i.

The next section introduces a distribution which is of interest in its own right.

3.6 Joint Limiting Distributions

It may be of interest to obtain the probability that a process arrives in state i, say, at time t and then, after a further period of time z, it is in state j; that is, Prob $\{ Z_t = i,\ Z_{t+z} = j \}$. In particular we would like to consider the case as $t \to \infty$. An example of a situation in which this may apply is if one wishes to find the steady state probability that at two instants of time z apart, the process is in states i and j, respectively. It should be pointed out that this probability says nothing about which states have been visited in between.

The main results of the section is summarized by

Theorem 2. If $H_{ij}(t)$ is continuous in t for all i,j, then

$$\text{Prob}\{Z_t = i,\ Z_{t+z} = j\} \to R_{ij}(z) \qquad \text{as } t \to \infty$$

where

$$R_{ij}(z) = \frac{\eta_i}{\mu_i}\left\{\sum_{k=1}^{N}\int_0^z (p_{ik}-F_{ik}(u))\,\phi_{kj}(z-u)\,du + \delta_{ij}\int_z^\infty (1-W_i(u))\,du\right\} \tag{29}$$

Proof : From first principles and for all i,j,

$$\text{Prob}\{Z_t=i,\ Z_{t+z}=j\} = \sum_{k=1}^{N}\int_{u<z} \text{Prob}\{Z_{t+z}=j \mid Z_t=i, Z_t^+ = k, V_t^+=u\} \times d\text{Prob}\{Z_t=i,\ Z_t^+=k, V_t^+\le u\} \tag{30}$$

We now consider the probability terms under the integral sign separately as $t \to \infty$. By the Markov property it follows that, for $u \le z$,

$$\text{Prob}\{Z_{t+z}=j \mid Z_t = i,\ Z_t^+ = k,\ V_t^+ = u\} = \phi_{kj}(z-u) \tag{31}$$

for all $t \ge o$. Also from (27) we note that

$$\text{Prob}\{Z_t = i,\ Z_t^+ = k,\ V_t^+ \le u\} = \sum_{m=1}^{N} a_m P_m\{Z_t=i,\ Z_t^+=k,\ V_t^+ \le u\}$$

where, as before, $a_m = \text{Prob}\{X_o = m\}$. Hence

$$\lim_{t\to\infty} \text{Prob}\{Z_t = i, Z_t^+ = k,\ V_t^+ \le u\} = \frac{\eta_i}{\mu_i}\int_o^u (p_{ik}-F_{ik}(y))\,dy \tag{32}$$

Since $\phi_{kj}(z-u)$ is a probability, it is bounded, and by virtue of the continuity of the $H_{ij}(\cdot)$, it is also continuous in $u(\le z)$ for all j,k and $z \ge o$. Consequently, the integral and

limit signs may be interchanged (Billingsley, 1979, pages 286 and 288). Thus (30) becomes

$$R_{ij}(z) = \lim_{t\to\infty} \sum_{k=1}^{N} \int_{u\le z} \phi_{kj}(z-u)\, dProb\{ Z_t = i,\ Z_t^+ = k,\ V_t^+ \le u \}$$

$$= \sum_{k=1}^{N} \int_{u\le z} \phi_{kj}(z-u)\, \frac{\eta_i}{\mu_i} (p - F_{ik}(u))\, du \qquad (33)$$

which, on rearranging, by the theorem of total probability yields the required result. That is

$$R_{ij}(z) = \sum_{k=1}^{N} \int_{u\le z} \phi_{kj}(z-u) \cdot \eta_i (p_{ik} - F_{ik}(u))\, du$$

$$+ \delta_{ij}\eta_i \sum_{k=1}^{N} \int_{z}^{\infty} (p_{ik} - F_{ik}(u))\, du$$

$$= \eta_i \{ \sum_{k=1}^{N} \int_{o}^{z} (p_{ik} - F_{ik}(u))\, \phi_{kj}(z-u)\, du$$

$$+ \delta_{ij} \int_{z}^{\infty} (1-W_i(u))\, du \}$$

If we sum (29) over j the result reduces to the limit η_i, as expected, irrespective of any initial state. Thus

$$\sum_{j=1}^{N} R_{ij}(z) = \frac{\eta_i}{\mu_i} \{ \sum_{k=1}^{N} \int_{o}^{z} (p_{ik} - F_{ik}(u)) \sum_{j=1}^{N} \phi_{kj}(z-u)\, du$$

$$+ \int_{z}^{\infty} (1 - W_i(u))\, du \}$$

$$= \frac{\eta_i}{\mu_i} \{ \int_{o}^{z} (1-W_i(u))du + \int_{z}^{\infty} (1-W_i(u))du \}$$

$$= \eta_i \qquad (34)$$

An analytically simpler form of $R_{ij}(z)$ is given by the relation

$$R_{ij}(z) = \frac{\eta_i}{\mu_i}(\delta_{ij}\mu_i + \int_o^z \left(\sum_{k=1}^N p_{ik}\,\phi_{kj}(u) - \phi_{ij}(u) \right) du) \qquad (35)$$

$$i,j = 1,2,\ldots, N.$$

To arrive at this result, first we note that, from (29)

$$\text{(i)} \quad \sum_{k=1}^N \int_o^z F_{ik}(u)\,\phi_{kj}(z-u)\,du = \sum_{k=1}^N \int_{o\le u\le z} \int_{v\le u} dF_{ik}(v)\phi(z-u)\,du$$

$$= \sum_{k=1}^N \int_{y\le z} \int_{v\le z-y} dF_{ik}(v)\,\phi_{kj}(z-y-v)\,dy \qquad (u=v+y)$$

$$= \int_o^z (\phi_{ij}(z-y) - \delta_{ij} \int_{z-y}^{\infty} dW_i(u))\,dy;$$

and

$$\text{(ii)} \quad \sum_{k=1}^N \int_o^z p_{ik}\phi_{kj}(z-u)\,du = \int_o^z \sum_{k=1}^N p_{ik}\phi_{kj}(z-u)\,du \qquad (36)$$

In addition, we may write

$$\text{(iii)} \quad \int_o^z (1-W_i(u))\,du = \sum_{k=1}^N \int_o^z (p_{ik} - F_{ik}(u))\,du$$

Substitutions of these results in (29) yields

$$R_{ij}(z) = \frac{\eta_i}{\mu_i} \int_o^z \left(\sum_{k=1}^N p_{ik}\phi_{kj}(u) - \phi_{ij}(u) \right) du$$

$$+\frac{\eta_i}{\mu_i}\delta_{ij} \int_z^{\infty} (1-W_i(u))\,du$$

$$+\frac{\eta_i}{\mu_i}\delta_{ij} \int_o^z (1-W_i(u))\,du$$

$$= \frac{\eta_i}{\mu_i}(\delta_{ij}\mu_i + \int_o^z \left(\sum_{k=1}^N p_{ik}\phi_{kj}(u) - \phi_{ij}(u) \right) du)$$

as required. $\qquad i,j = 1, 2,\ldots,N$

The Laplace transform of the system of equations (35) is given by

$$R^{*}_{ij}(\theta) = \frac{\eta_i}{\theta\mu_i}\left(\delta_{ij}\mu_i + \sum_{k=1}^{N} p_{ik}\phi^{*}_{kj}(\theta) - \phi^{*}_{ij}(\theta)\right) \qquad (37)$$

$$i,j = 1, 2, \ldots, N$$

3.7 An Illustrative Example

When the distribution functions of sojourn times between transitions are expressible as simple rational Laplace-Stieltjes transforms, reasonably simple results for (35) and therefore (29) are forthcoming. For illustration, we discuss the special case of a 2-state (N=2) stochastic process, which can also be handled, probably more elegantly, by other methods such as the method of renewal process.

The transition matrix $\underline{P}$, of the embedded Markov chain is of the form

$$\underline{P} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \qquad \begin{matrix} o \le a_{ij} \le 1 \\ i,j = 1, 2 \end{matrix} \qquad (38)$$

We further assume that the times between transitions are exponentially distributed, with parameters $\lambda_{ij} \ge o$, $i,j = 1,2$. As a consequence, the distribution functions $H_{ij}(\cdot)$ have the Laplace-Stieltjes transforms $H^{**}_{ij}(\theta) = \frac{\lambda_{ij}}{\theta+\lambda_{ij}}$,

from which we get

$$\underline{F}^{**}(\theta) = \begin{bmatrix} \frac{a_{11}\lambda_{11}}{\theta+\lambda_{11}} & \frac{a_{12}\lambda_{12}}{\theta+\lambda_{12}} \\ \frac{a_{21}\lambda_{21}}{\theta+\lambda_{21}} & \frac{a_{22}\lambda_{22}}{\theta+\lambda_{22}} \end{bmatrix} \qquad (39)$$

If we now substitute for $\underline{F}^*(\theta)$ in (9), noting that $\underline{F}^{**}(\theta) = \theta\underline{F}^*(\theta) - \underline{F}(o)$ ($\underline{F}(o) = \underline{O}$)), then we obtain

$$\phi^*_{ij}(\theta) = \begin{cases} a_{ij}\left(\dfrac{\theta+a_{ji}\lambda_{jj}}{\theta+\lambda_{ii}} + \dfrac{a_{jj}\lambda_{ij}}{\theta+\lambda_{jj}}\right) \Big/ (\mathrm{Det}(\theta)(\theta+\lambda_{ij})) \\ \qquad\qquad i \neq j \\ \left(\dfrac{a_{ii}(\theta+a_{3-i}\lambda_{3-i\ 3-i})}{(\theta+\lambda_{ii})(\theta+\lambda_{3-i\ 3-i})} + \dfrac{a_{i3-i}\,a_{3-ii}\lambda_{i3-i}}{(\theta+\lambda_{i\ 3-i})(\theta+\lambda_{3-ii})}\right) \Big/ \mathrm{Det}\ (\theta) \\ \qquad\qquad i = j, \end{cases} \tag{40}$$

where Det (θ) is the determinant of $(\underline{I}-\theta\underline{F}^*(\theta))$. Explicit expressions for the $R_{ij}(\cdot)$ are then obtained by substituting (40) in (37) for the $\phi^*_{ij}(\theta)$ and extracting the inverse Laplace transforms of the result.

Suppose that $a_{11} = a_{22} = o$, in which case $\lambda_{11} = \lambda_{22} = o$, then the process reduces to an alternating Poisson process and may also be analysed as such. In our example, (40) now becomes in matrix form

$$\underline{\phi}^*(\theta) = \begin{bmatrix} \dfrac{\lambda_{12}}{\theta(\theta+\lambda_{12}+\lambda_{21})} & \dfrac{\theta+\lambda_{21}}{\theta(\theta+\lambda_{12}+\lambda_{21})} \\ \dfrac{\theta+\lambda_{12}}{\theta(\theta+\lambda_{12}+\lambda_{21})} & \dfrac{\lambda_{21}}{\theta(\theta+\lambda_{12}+\lambda_{21})} \end{bmatrix} \tag{41}$$

From (37) $\underline{R}^*(\theta)$ is found to be

$$\underline{R}^*(\theta) = \frac{1}{\theta(\theta+\lambda_{12}+\lambda_{21})} \begin{bmatrix} \eta_1(\theta+\lambda_{12}) & \eta_1\lambda_{21} \\ \eta_2\lambda_{12} & \eta_2(\theta+\lambda_{21}) \end{bmatrix} \tag{42}$$

From (40) the inverse Laplace transforms are easily obtained.

$$\underline{R}(Z) = \begin{bmatrix} \eta_1 \quad (\eta_1 + \eta_2 \quad g(z)) & \eta_1\eta_2 \quad (1-g(z)) \\ \eta_1\eta_2 \quad (1-g(z) & \eta_2 \quad (\eta_2 + \eta_1 \quad g(z)) \end{bmatrix} \tag{43}$$

where

$$g(z) = \exp(-z(\lambda_{12} + \lambda_{21})) \tag{44}$$

3.8 Discussion

Like most stochastic models, semi-Markov models do have some serious limitations as well as advantages.

First, finding the transition probability distributions, $\phi_{ij}(\cdot)$, is mathematically intractable, except for the most simple forms of $H_{ij}(\cdot)$. This is clearly shown in the example in the last section. One way to proceed would seem to be by simulation, a procedure to be pursued in Chapter 5, or by numerical integration, as in Chapter 6.

A second difficulty with these models is that it is not always possible to obtain data, from the processes being modelled, in the level of detail that would allow adequate model validation. Moreover, these models usually contain many parameters for which only ad-hoc ranges of values may be used.

An important advantage of semi-Markov models is perhaps that they are quite flexible and can accommodate rather complex situations. They can, for instance, be constructed such that they incorporate the non-random features of, and the causal structures underlying the process being modelled (Ginsberg 1971 b,c). However, this advantage is diminished to some extent by the arbitrarily limited dependence of sojourn times on jump structure.

In the next chapter, we propose and describe a semi-Markov model of malaria, and suggest some forms for the probability distribution functions for the sojourn times in each state.

CHAPTER FOUR

MODELLING THE TIME DISTRIBUTIONS

4.1 Basic assumptions

To recapitulate from Chapter 1, malaria transmission in a community involves two different host populations and several life-cycles of the parasite population, in addition to a host of other socio-environmental factors. As a consequence, any mathematical model of the course of the disease that attempts at embracing these characteristics explicitly, is likely to be intricate, impractical and may lead to intractable mathematics.

One way to obtain a reasonable mathematical statement of the course of the disease may be by constructing a hybrid semi-Markov model. Such a model would be made to incorporate, for example, some of the vector effects, deterministically while treating as random the times spent by an individual in some of the phases of the disease. It may be argued that the vector population is usually large relative to the human population, so that the inclusion of its effects as deterministic would be both appealing and appropriate. An example of combining deterministic variables with random variables is that of Nasell and Hirsh (1973) and Nasell (1976a,b, 1977), in modelling schistosomiasis (Bilharzia) in an isolated community. In their rather intricate model, the snail population, analogous here to the mosquito population, is included as deterministic in an otherwise stochastic model.

We shall consider an idealized situation of a relatively isolated community in an endemic malaria environment.

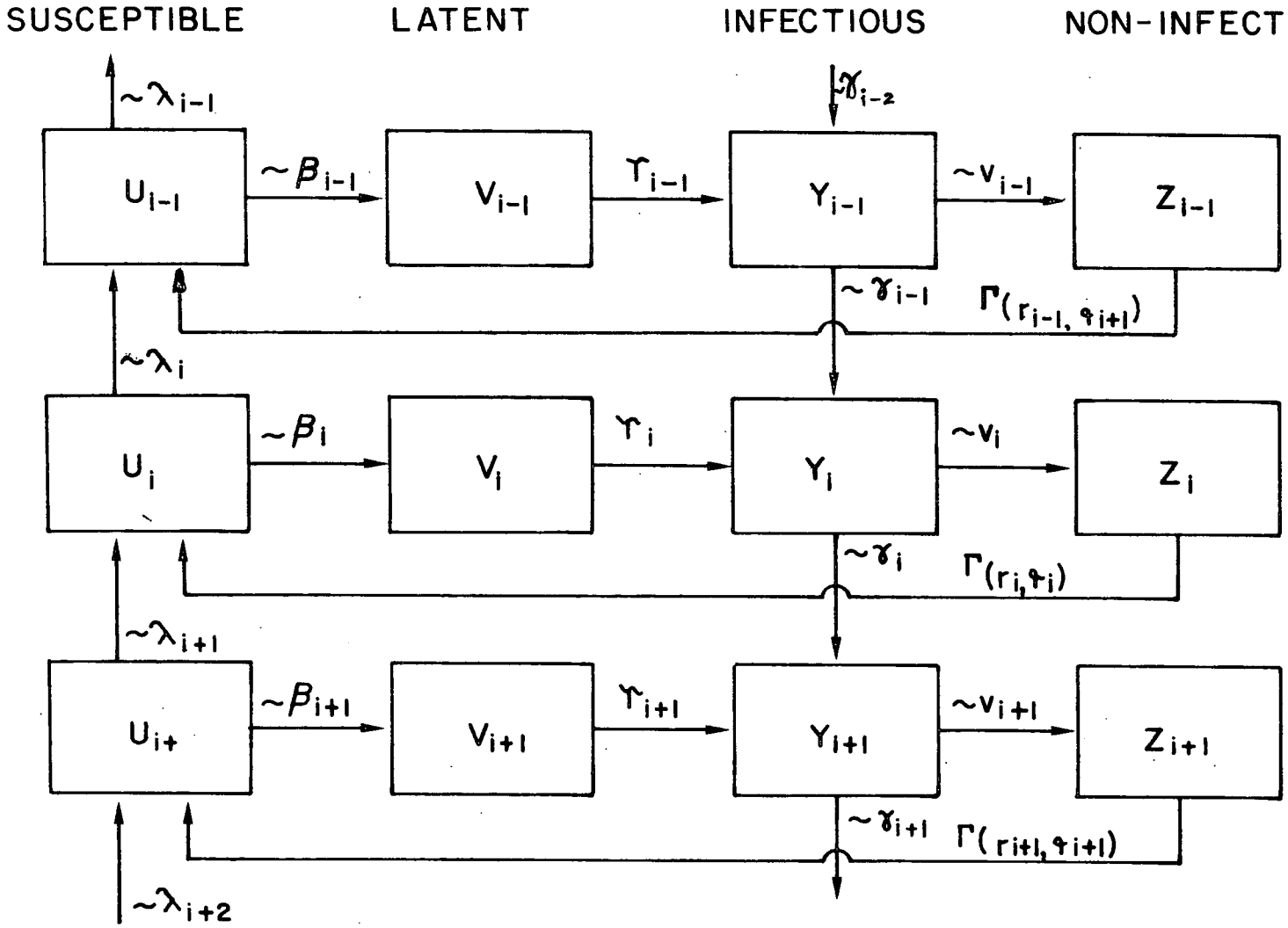
In addition, we shall assume that the community has an approximately constant population, whose members are not subject to birth, death or migration. Let us also suppose that each member of the community belongs to one of four basic categories with respect to malaria infection. These categories are (i) the susceptible, (ii) latent, (iii) infectious, and (iv) the non-infectious (but infected). The sets of the individuals in each of the categories will be denoted by U, V, Y and Z, respectively.

As pointed out earlier (Chapter 1, section 1.5), both the acquisition and loss of acquired immunity are gradual. But, for convenience, we shall treat the degree of immunity in an individual as discrete, such that persons who enjoy high measure of immunity are classified in the higher immunity levels, with the non-immunes constituting the lowest level. The sets of individuals are accordingly sub-divided, each into a finite number of subsets, U_i, V_i, Y_i and Z_i, $i = 1,2,\ldots,k$, where k is the number of immunity levels.

4.2 Model Description

The model is best described by following the phases through which one individual passes before and during an infection. As we proceed through the phases, we shall sometimes try and interpret empirical evidence to make certain assumptions about, for example, distributions of sojourn times in these phases. We begin by assuming that the person has only just arrived into the susceptible category at immunity level i; that is, he is in set U_i at time $t = o$, but was not there at $t = o-$. Figure 8 gives a schematic diagram of the model.

Fig. 8 DIAGRAMMATIC REPRESENTATION OF THE STATES AND TRANSITIONS OF THE SEMI-MARKOV MALARIA MODEL



Except for the τ_i, the other perameters are of the transition time distributions as described in text. τ_i are the constant times to be spent in the latent states before transition is made to the infectious states.

4.2.1 The Susceptible Stage

In this stage the person is free of the disease but is continuously exposed to the risk of infection if bitten by an infectious mosquito. As time goes by, it is assumed one of two situations may occur independently of each other. The person either becomes infected and moves to the latent set V_i, or he loses his current immunity level and slips into the next lower level in the susceptible category, U_{i-1} $(1 < i \leq k)$. Thus, his rates of transitions to V_i and U_{i-1} are given respectively by:

$$\text{(i)} \quad \text{Prob}\{U_i \to V_i \text{ in } (t,t+\delta t)\} = \beta_i(t)\delta t + o(\delta t)$$

$$\text{(ii)} \quad \text{Prob}\{U_i \to U_{i-1} \text{ in } (t,t+\delta t)\} = \lambda_i \delta t + o(\delta t) \qquad (1)$$

where $\beta_i(t)$ is the infection rate of the person at time t, and $\lambda_i > o$ is the rate of loss of immunity. For $i = 1$, $\lambda_i = o$. In addition, since acquired immunity is developed only during an infection (Section 1.5), no transition to a higher immunity level is expected to take place in this stage. In the next section (4.3), we suggest and discuss a specific form for $\beta_i(t)$.

4.2.2 The Latent Stage

This phase covers the period between the time sporozoites are introduced into the blood stream and the time the infected person becomes infectious.

There are wide variations in the period, depending on the particular strain or species of the parasite (section 1.1). For P.falciparum, it is about 17 to 21 days. This estimate

may be inferred from the following. First, we may recall (section 1.3) that the pre-erythrocytic phase is estimated to last from $5\frac{1}{2}$ to 7 days. Secondly, the sexual forms of the parasite appear about 10 days after the start of the erythrocytic phase. And finally, other investigations (Jeffery et al., 1955) have shown that gametocytes are functionally mature only after about 2 to 4 days of the start of gametogony. We shall assume as a first approximation, that latent periods, denoted by τ_i, are constant for each immunity level i.

4.2.3 The Infectious Stage

Mature and infective gametocytes are now present in the blood stream of the host. This is clearly the most crucial phase of the disease as far as its perpetuation in a community is concerned. Some of the factors that may be expected to influence the infectivity of gametocytes, and hence the infectious stage, include their longevity, density, sex ratio and any immune response of their host that may either limit their infectivity or persistence (Smalley et al., 1976). We begin by considering the immunity factor.

Studies about the existence and nature of any immune response to P. falciparum gametocytes are inconclusive (ibid.), but as we mentioned earlier (section 1.5), it is currently accepted that immune response has little or no effect on gametocytes. However, it has also been shown (McGregor et al., 1965; Molineaux et al., 1980, pages 115-118) that gametocyte density appears to fall as the age of the host advances, and hence, with rising degree of immunity. This is usually

interpreted to mean that individuals with high degree of immunity are less infectious than those with low or no immune responses (Dietz et al., 1974, Sambasivan, 1975). As regards superinfection, it is known to have little or no effect on gametocyte production; see for instance Dietz et al. It should therefore have negligible effect on the distribution time in this stage.

Smalley et al. also found that gametocyte production in a host usually rose to a peak, after which the gametocyte population would begin to fall at approximately constant rate. Furthermore, they observed that the ratio of the female to male gametes was 4 to 1 and was maintained throughout up to the depletion of their numbers. Thus both sexual forms were equally likely to be removed from the peripheral blood at all times.

No mention was made about when gametocyte production attained its peak from the time gametogony set in. We shall, however, assume that the time lapse is small relative to the mean time gametogony lasts. Hence, with the rate of loss of gametocytes taken as constant, it is postulated that infectious times are exponentially distributed with mean $\frac{1}{\nu_i}$ and $\nu_i \leq \nu_{i+1}$.

We now consider the problem of gain in immunity. Tobie et al. (1966a,b) conducted studies to demonstrate the production and persistence of antibody response in a group of volunteers infected with P.vivax. They observed that antibody levels rose abruptly about 6 days after the onset of parasitamia to plateau some 8 days later. After about 3 weeks they gradually

declined to low but sustained levels, following the elimination of the parasites. Presumably, the low levels were higher than those before infection which would allow for the slow build-up of these levels over the years.

Similar studies, carried out earlier by Collins et al. (1964) for P.falciparum infections, were inconclusive because of the few number (3) of the subjects involved. But their results show strong similarity, and it is possible that the situations for the two species may not be that different. We shall therefore assume that immune responses, for P. falciparum infection, start building up towards the end of the latent stage and attain maxima some time during the infectious stage.

Thus we shall suppose that the gain in the level of immunity in the model takes place during the infectious stage as shown in Figure 8. Furthermore, it will be assumed that the boost in antibody response may be approximated by a Poisson process of rate γ_i. Hence, given that a person is in subset Y_i for time T, he may gain immunity and move to a higher level Y_{i+1} with probability $1-e^{-\gamma_i T}$. Otherwise, if the immune response boost is below some threshold, he remains in the same level of immunity (with probability $e^{-\gamma_i T}$), in which case his next move is to Z_i.

4.2.4 The Non-infectious stage

Here, while the individual still exhibits the presence of parasites in his blood, he no longer has any clinically detectable gametocytes. On recovery the person moves into the susceptible stage, still retaining his immunity level.

Unlike the previous stage, the effect of superinfection

here is important. As this phenomenon has been discussed at some length in section 3 of Chapter 2, suffice to recall here that superinfection is supposed to prolong the presence of parasites in a host. Moreover, the longer the person stays in this state the more he is at risk of re-infection, thus suggesting a time distribution with decreasing hazard function. Because of its simplicity and flexibility, we propose a gamma distribution of the non-infectious time, with parameters r_i and α_i, where $o < r_i < 1$ for all i, and $\alpha_i/r_i \leq \alpha_{i+1}/r_{i+1}$, $i = 1,2,\ldots, k-1$. The last inequality is because persons in higher immunity levels have faster recovery rates (section 1.5).

In the next section we dwell a bit further on the distribution of the time to infection.

4.3 On the Distribution of the Time to Infection

The occurrence of successful infectious contacts on a susceptible may be thought of as a time-dependent Poisson process. Thus, if $\beta_i(\cdot)$ is the hazard function of the distribution, the probability that at least one such contact is made on a person before time t+v, given that he entered U_i at time t, is given by (cf. equation (28) of Chapter 2).

$$H_i(t,v) = 1 - \exp(-\Lambda_i(t,v)) \qquad (2)$$

where

$$\Lambda_i(t,v) = \int_t^{t+v} \beta_i(u)du \qquad (3)$$

It would seem unrealistic to express $\beta_i(\cdot)$ as a single constant, except perhaps for relatively small time periods.

For instance, $\beta_i(\cdot)$ should depend in some sense on either the vector population, the factors influencing it, the vector's propensity to transmit the disease or on all of these. However, as already stated (Chapter 1), these factors interact in complex ways some of which are still not well understood. To incorporate their effect explicitly is highly impractical. In an attempt to overcome this difficulty, we shall prescribe a rather simple and intuitive form for $\beta_i(\cdot)$ which is aimed at reflecting only the most salient features of these factors.

Mention has already been made, in Chapter 1, of how malaria prevalence tends to fluctuate markedly with the season in an endemic community. Usually, the disease prevalence is highest at the peak of the wet season and least during the dry season, coinciding with the rise (wet season) and fall (dry season) in the density of the vector population. Hence and for the sake of simplicity a particular form for $\beta_i(\cdot)$ is proposed which contains a term varying sinusoidally with time in addition to a constant term. The sinusoidal term is assumed to have a period of one year, the seasonal cycle. Thus, for all i,

$$\beta_i(t) = b_i(1-b_i' \cos \omega t), \quad 0 < b_i' < 1 \qquad (4)$$

where $\omega = 2\Pi/365$ (365 being the number of days in one calendar year), or some other appropriate frequency value, and b_i' is the relative amplitude of the sinusoidal modulation. Without loss of generality, the origin of time t is taken to be the beginning of one 'epidemiological' year, the period between two consecutive minimum levels of disease prevalence, assumed equal to that of a calendar year.

If we substitute for $\beta_i(\cdot)$ in (3), dropping the subscript i for the moment, we get

$$\Lambda(t,v) = b \int_t^{t+v} (1-b' \cos\omega u)\, du$$

$$= b \{ v - \frac{b'}{\omega} (\sin\omega(t+v) - \sin\omega t) \} , \qquad (5)$$

so that,

$$H(t,v) = 1 - \exp \{ -b(v - \frac{b'}{\omega} (\sin\omega(t+v) - \sin\omega t)) \} , \qquad (6)$$

which reduces to the usual exponential distribution when b' = o.

This form (4) of $\beta(\cdot)$ has been used before to investigate the seasonality effect of the prevalence of seasonally varying disease in a strictly deterministic sense. Dietz (1976) has used the form of $\beta(\cdot)$ suggested here as the infection rate and applied it in a deterministic model. An analogous form of $\beta(\cdot)$ is used by Bailey (1975, page 138) also as an infection rate in a deterministic model. We may recall (Chapter 2, sub-section 2.2.2) that Aron and May (1980) have suggested modifying the basic Ross-Macdonald model by introducing a variable vectorial density of the form (4). In a different context, that of the study of cosmic rays, Willis (1964) has discussed some properties of a distribution whose hazard function is essentially of the form (4).

Let $\mu_1(\cdot)$ and $\mu_2(\cdot)$ denote the first and second moments respectively of the distribution defined by (6). Then

$$\mu_1(t) = \int_o^\infty v\, d_v H(t,v)$$

$$= b \int_o^\infty v\, (1-b' \cos\omega(t+v))$$

$$x \exp\left\{ -b\left(v - \frac{b'}{\omega}(\sin\omega(t+v) - \sin\omega t)\right)\right\} dv$$

$$= b \int_0^\infty v\,(1-b'\cos\omega(t+v))\exp(-bv)$$

$$x \sum_{j=0}^{\infty} \frac{\left(\frac{bb'}{\omega}\right)^j \left(\sin\omega(t+v)-\sin\omega t\right)^j}{j!} dv \qquad (7)$$

where $\exp\left\{\frac{bb'}{\omega}(\sin\omega(t+v) - \sin\omega t)\right\}$ is replaced by its series form. Although the integral seems unmanageable, an approximate solution can be obtained when $bb' << 1$. In this case we drop all except the first two terms of the series expansion. Equation (7) then becomes

$$\mu_1(t) \simeq b \int_0^\infty v\,(1-b'\cos\omega(t+v))\exp(-bv)$$

$$x \left\{ 1 + \frac{bb'}{\omega}(\sin\omega(t+v) - \sin\omega t)\right\} dv$$

$$= b\left(1 - \frac{bb'}{\omega}\sin^{\omega}t\right) \int_0^\infty v \exp(-bv)\, dv$$

$$+ \frac{bb'}{\omega}(b\sin\omega t - \omega\cos\omega t) \int_0^\infty v\cos\omega v \exp(-bv)dv$$

$$+ \frac{bb'}{\omega}(\omega\sin\omega t + b\cos\omega t) \int_0^\infty v \sin\omega v \exp(-bv)\, dv \qquad (8)$$

On performing the integration (Gradshteyn and Ryzhik, 1980, page 198), and simplifying the result, we obtain

$$\mu_1(t) \simeq \frac{1}{b} + c_1\sin\omega t + c_2\cos\omega t \qquad (9)$$

where

$$(i) \quad c_1 = -\frac{b'\omega}{(b^2+\omega^2)^2}(5b^2+\omega^2)$$

$$(ii) \quad c_2 = -\frac{bb'}{(b^2+\omega^2)^2}(3b^2-\omega^2)$$

Similarly, the second moment, $\mu_2(t)$ is obtained as

$$\mu_2(t) \simeq \frac{2}{b^2} + d_1 \sin\omega t + d_2 \cos\omega t \quad , \tag{10}$$

where

$$\text{(i)} \quad d_1 = \frac{2bb'}{\omega} \left(\frac{(b-\omega)^4}{(b^2+\omega^2)^3} - \frac{1}{b^2} \right)$$

$$\text{(ii)} \quad d_2 = \frac{8b^2b'(b^2-\omega^2)}{(b^2+\omega^2)^3}$$

The approximate variance $\sigma^2(t)$, is then given by

$$\sigma^2(t) \simeq \frac{1}{b^2} + \left(d_1 - \frac{2c_1}{b}\right) \sin\omega t + \left(d_2 - \frac{2c_2}{b}\right)\cos\omega t \tag{11}$$

Thus both the mean and variance of the distribution are also sinusoidal. When $b' = o$ then $c_i = d_i = o$ for $i = 1,2$, and we then obtain the familiar versions for the exponential distribution.

4.4 Some General Comments

In this chapter, we have tried to construct models for the distribution of times between transitions, based on some features underlying the process, coupled with mathematical convenience. The rather ad hoc nature of the latter is partly offset by the choice, in some cases, of a flexible family of distributions, namely the gamma.

The classification of individuals into discrete categories of susceptibles, latent, infectious, non-infectious may not be realistic. Mention has been made of the fact that parasites usually exist at various densities within their hosts, and that the epidemiology of malaria depends more on these quantitative aspects of infection than on mere presence and absence of parasites (Section 2.4). Moreover, the acquisition and loss of immunity are gradual processes.

We also note that, in postulating the hazard function of the distribution time to infection, it would have seemed more realistic if, for instance, vector density and proportion of infectious persons in the community, were explicitly included. However, this would have entailed intractable mathematics. We expect that the expression suggested reflects the combined effects of these factors.

In spite of the shortcomings mentioned above and certain omissions (such as birth, death and migration), it is hoped that the model described here portrays some of the more salient characteristics of the disease and its persistence in an endemic environment, in a way capable of providing sensible conclusions. In the next chapter, we discuss the application of the model and its simulated version.

CHAPTER FIVE

A SEMI-MARKOV MODEL OF MALARIA

5.1 Introduction

Having set up a semi-Markov framework to describe the various stages of malaria infection, we now consider the model in greater detail. In particular we shall discuss, for several values of k (the number of immunity levels) the steady-state solutions together with the results of varying some key parameters. The more complex cases are investigated by simulation procedures. We conclude by a discussion of some examples of simulation runs and the deficiencies of the model.

5.2 The steady-state model

We begin by considering the case in which seasonality has negligible effect on infection rates. This case should be approximately correct over short periods relative to seasonal variations. With $b_i' = o$, $i = 1,2,\ldots, k$, in equation (4) of Chapter 4, and in the absence of change of immunity, times to infection are now exponentially distributed with means $1/\beta_i$.

The transition probability matrix $\underline{P}$, of the embedded Markov chain of the process may be written as

$$\underline{P} = \begin{bmatrix} \underline{B} & \underline{C} & \underline{0} & \underline{0} \\ \underline{0} & \underline{0} & \underline{I} & \underline{0} \\ \underline{0} & \underline{0} & \underline{D} & \underline{E} \\ \underline{I} & \underline{0} & \underline{0} & \underline{0} \end{bmatrix} \qquad (1)$$

where $\underline{B}$, $\underline{C}$, $\underline{D}$ and $\underline{E}$ are k x k submatrices of the simple forms

$$\underline{B} = \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ b_2 & 0 & \cdots & 0 & 0 \\ 0 & b_3 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & b_k & 0 \end{bmatrix}$$

$$\underline{C} = \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1-b_2 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 1-b_k \end{bmatrix}$$

$$\underline{D} = \begin{bmatrix} 0 & d_1 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & d_{k-1} \\ 0 & 0 & \cdots & 0 & 0 \end{bmatrix}$$

$$\underline{E} = \begin{bmatrix} 1-d_1 & 0 & & 0 & 0 \\ 0 & 1-d_2 & & 0 & 0 \\ & & & & \\ 0 & 0 & & 1-d_{k-1} & 0 \\ 0 & 0 & & & 1 \end{bmatrix}$$

$\underline{I}$ and $\underline{0}$ are k x k identity and null matrices, respectively.

All the states in the model inter-communicate and, therefore, the embedded Markov chain is irreducible. Furthermore, if the chain is aperiodic then there is a unique row vector $\underline{\Pi}$ satisfying the system of equations (17) in Chapter 3, which is

$$\underline{\Pi}\,\underline{P} = \underline{\Pi}$$

$$\sum_{\ell=1}^{m} \Pi_\ell = 1, \quad \Pi_j \geq 0 \quad , \; j = 1,2,\ldots, m,$$

where m = 4k, is the number of states of the process.

Let $\underline{\Pi}_u$, $\underline{\Pi}_v$, $\underline{\Pi}_y$ and $\underline{\Pi}_z$ denote the 1 x k row vectors whose components are those of $\underline{\Pi}$ corresponding to the subsets of U, V, Y and Z, respectively. Then, substituting (1) for $\underline{P}$ in the vector equation $\underline{\Pi}\underline{P} = \underline{\Pi}$, we obtain

$$\begin{aligned}
&\text{(i)} \quad \underline{\Pi}_u = \Pi_u \underline{B} + \underline{\Pi}_z \; ; \\
&\text{(ii)} \quad \underline{\Pi}_v = \underline{\Pi}_u \underline{C} \; ; \\
&\text{(iii)} \quad \underline{\Pi}_y = \underline{\Pi}_v + \underline{\Pi}_y \underline{D} \; ; \\
&\text{(iv)} \quad \underline{\Pi}_z = \underline{\Pi}_y \underline{E} \; .
\end{aligned} \tag{2}$$

Thus, with the dimension reduced by blocking, it now becomes quite simple to solve for $\underline{\Pi}$. The unconditional equilibrium probabilities, η_j, of being in any state j are then obtained by substituting for the Π_i in equation (21) of Chapter 3.

We now consider the specific cases of k = 1,2. The analysis for k $\geq$ 3 may be carried out in similar fashion although with increasing level of cumbersome but otherwise straightforward algebra.

5.3 Two specific examples

(a) Case k = 1. In this rather trivial situation, it is assumed that an individual's immune response is either inoperative or that it remains unaltered by fresh attacks or recoveries.

The process now reduces to a 4-state process, with the transition matrix (1) given by

$$\underline{P} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix} \qquad (3)$$

Clearly (3) defines a process that moves deterministically from one state to the other. Alternatively, if $\underline{P}$ is thought of as a (degenerate) doubly stochastic matrix, then all states of the embedded Markov chain have equal chance of being occupied in the limit with $\underline{\Pi} = (\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4})$.

From the distributions of sojourn times described in the last chapter, and noting b' = o, the expected durations μ_j, j = 1, 2, 3, 4, are $\frac{1}{\beta}$, τ, $\frac{1}{\nu}$, $\frac{r}{\alpha}$, for the susceptible, latent, infectious and non-infectious states (stages), respectively. If we now substitute for Π_j and μ_j in equation (21) of Chapter 3, the unconditional probabilities, η_i, are

given by

$$\text{(i)}\quad 1/\{1+\beta(\tau+\frac{1}{\nu}+\frac{r}{\alpha})\};\quad \text{(the susceptible state)}$$

$$\text{(ii)}\quad \beta\tau/\{1+\beta(\tau+\frac{1}{\nu}+\frac{r}{\alpha})\};\quad \text{(the latent state)}$$

$$\text{(iii)}\quad \beta/\{\nu(1+\beta(\tau+\frac{1}{\nu}+\frac{r}{\alpha}))\};\quad \text{(the infectious state)} \qquad (4)$$

$$\text{(iv)}\quad \beta r/\{\alpha(1+\beta(\tau+\frac{1}{\nu}+\frac{r}{\alpha}))\};\quad \text{(the non-infectious state).}$$

A rough idea of the result of varying some of the parameters may be gained from these equations. For instance, the (unconditional) probability of occupying the susceptible state, η_1, is greater when the rate of infection β is reduced by a factor of 2 than when the recovery rate $\frac{\alpha}{r}$ is doubled. Thus, from (4 (i)) we have

$$\frac{2}{2+\beta(\tau+\frac{1}{\nu}+\frac{r}{\alpha})} - \frac{1}{1+\beta(\tau+\frac{1}{\nu}+\frac{r}{2\alpha})}$$

$$= \frac{\beta(\tau+\frac{1}{\nu})}{(2+\beta(\tau+\frac{1}{\nu}+\frac{r}{\alpha}))(1+\beta(\tau+\frac{1}{\nu}+\frac{r}{2\alpha}))}$$

$$> 0 \qquad (5)$$

This inequality may be interpreted to mean the following. From the public health viewpoint, and all things being equal (such as, for instance, the cost per proportionate change in parameters, facilities and expertise), the policy of using insecticides aimed at a reduction in β is preferable to that

of, say, the application of mass drug administration, which would bring about an increase in α/r, the rate of recovery. However, we have ignored here the fact that a reduction in infection rate would also reduce superinfection and so accelerate recovery. In this case the inequality (5) may not necessarily hold. Furthermore, although some drugs such as chloroquine are known to have no effect on the infectivity of mature gametocytes (Smalley et al, 1977) most malaria drugs do. So the application of mass drug administration would also mean a reduction in the number of infectious individuals and therefore that of infected mosquitoes. The result would be a fall in β, the infection rate. In view of these remarks, inequality (5) may be considered only as illustrative.

(b) $k = 2$. Here the human population is divided only into two groups, the immune and the non-immune. This model may be thought of as a stochastic analogue of the modified model of Dietz et al. in section 2.4. One difference to be noted is that, here, we allow gain in immunity levels to occur between adjacent infectious states (Figure 9), following what has been stated in the last chapter (Section 4.2). We further assume here that the latent times are both equal to τ.

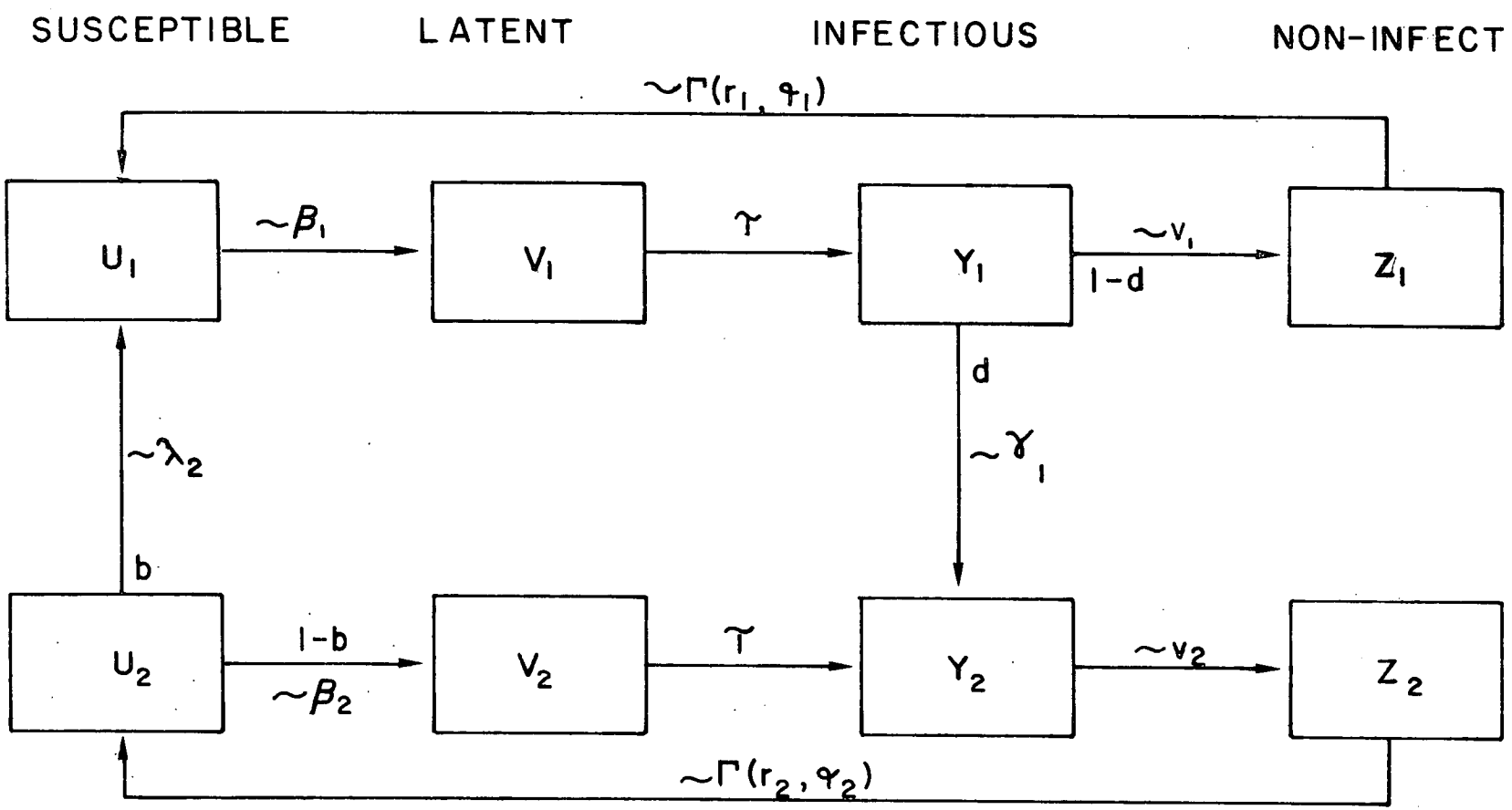
We have 8 states with the elements of $\underline{P}$ given by the following submatrices (from equation (1)).

$$\underline{B} = \begin{bmatrix} 0 & 0 \\ b & 0 \end{bmatrix}, \quad \underline{C} = \begin{bmatrix} 1 & 0 \\ 0 & 1-b \end{bmatrix}, \quad \underline{D} = \begin{bmatrix} 0 & d \\ 0 & 0 \end{bmatrix},$$

$$\underline{E} = \begin{bmatrix} 1-d & 0 \\ 0 & 1 \end{bmatrix} \qquad (6)$$

where $0 \leq b \leq 1$ and $0 \leq d \leq 1$. Substituting for

Fig. 9 SEMI-MARKOV TRANSMISSION MODEL OF MALARIA WITH TWO IMMUNE CATEGORIES



$\underline{B}$, $\underline{C}$, $\underline{D}$, and $\underline{E}$ in (2) and solving the result, yield

$$\underline{\Pi} = \{ \Pi_1, \Pi_1 d/b, \Pi_1, \Pi_1 d(1-b)/b, \Pi_1, \Pi_1 d/b, \Pi_1 (1-d), \Pi_1 d/b \} \quad (7)$$

where the probabilities are in the order of the subsets U_1, U_2, V_1, V_2, Y_1, Y_2, Z_1, Z_2, and

$$\Pi_1 = \frac{1}{4} \quad (1 + d/b - d/2)^{-1} \quad (8)$$

We note that (7) can also be got directly from Figure 9 fairly quickly.

If we define by $\underline{\psi}$ the row vector of the mean sojourn times in the 8 states, then

$$\underline{\psi} = \{ \frac{1}{\beta}, \quad \frac{1}{\beta_2 + \lambda_2}, \quad \tau, \quad \tau, \quad \frac{1-d}{\nu_1} + \frac{d}{\gamma_1}, \quad \frac{1}{\nu_2}, \quad \frac{r_1}{\alpha_1}, \quad \frac{r_2}{\alpha_2} \} \quad (9)$$

From (7) and (9) we then obtain, as before, the limiting probabilities of being in any state of the disease. We write below only the probabilities of being in the susceptible states (1 and 2). Those of the other states may be similarly derived.

$$\begin{aligned} &\text{(i)} \quad \eta_1 = \Pi_1 / (\beta_1 \eta_o) \\ &\text{(ii)} \quad \eta_2 = \Pi_1 d / \{ b (\beta_2 + \lambda_2) \eta_o \} \end{aligned} \quad (10)$$

where

$$\eta_o = \sum_{\ell=1}^{m} \Pi_\ell \mu_\ell$$

$$= \Pi_1 \{ \frac{1}{\beta_1} + \frac{d}{b(\beta_2 + \lambda_2)} + \tau(1 + \frac{d}{b} - d) + \frac{1-d}{\nu_1} + \frac{d}{\gamma_1} + \frac{d}{b\nu_2} + \frac{(1-d)r_1}{\alpha_1} + \frac{dr_2}{b\alpha_2} \} \quad (11)$$

As in the case k = 1, suppose we would like to use the model to choose some form of control measure. One of the criteria could be to choose one that would change the parameters affecting the largest increase in the value of say η_1 and η_2 or their sum. For convenience we shall denote $\eta_1 + \eta_2$ by η.

Let us consider for instance the case d = 1. This is the situation in which we assume that every infected individual from the non-immune level always gains immunity. This is probably a fairly reasonable assumption, considering that immune response is boosted every time one is infected. η is then expressed as,

$$\eta = \Pi_1 \left(\frac{1}{\beta_1} + \frac{1}{\lambda_2}\right)/\eta_o \quad , \tag{12}$$

where the transition probability, b, is now replaced by $\lambda_2/(\beta_2 + \lambda_2)$, and Π_1 and η_o are of the forms

$$\text{(i)} \quad \Pi_1 = \lambda_2 /(4\beta_2 + 6\lambda_2)$$

$$\text{(ii)} \quad \eta_o = \Pi_1 \left(\frac{1}{\beta_1} + \frac{1}{\lambda_2} + \frac{1}{\gamma_1} + \frac{(\beta_2 + \lambda_2)}{\lambda_2}\left(\tau + \frac{1}{\nu_2} + \frac{r}{\alpha_2}\right)\right) \tag{13}$$

Using the notation of Dietz et al. (1974), let $R_i = \frac{\alpha_i}{r_i}$, i = 1, 2, be the rate of transfer from the non-infectious state to the susceptible state. We shall choose the β_i and the R_i as our typical parameters of interest for control purpose. If we assume that all the parameters are independent of each other, then it is easily shown that the partial derivatives of (12) are negative with respect to the β_i and positive with respect to R_2. The signs indicate that η increases as either the β_i decrease or R_2 increases.

Informally, we can compare the effect of η by varying either of the parameters and examining the ratio of the appropriate absolute partial derivatives of η.

Thus, for β_2 and R_2 for example we have,

$$\text{(i)} \quad \frac{\partial \eta}{\partial \beta_2} = -\Pi_1^2 \left(\frac{1}{\beta_1} + \frac{1}{\lambda_2}\right)\left(\tau + \frac{1}{\nu_2} + \frac{1}{R_2}\right)/(\lambda_2 \eta_o^2)$$

$$\text{(ii)} \quad \frac{\partial \eta}{\partial R_2} = \Pi_1^2 \left(\frac{1}{\beta_1} + \frac{1}{\lambda_2}\right)(\beta_2 + \lambda_2)/(\lambda_2 R_2 \eta_o^2)$$

so that

$$\left|\frac{\partial \eta}{\partial \beta_2}\right| \Big/ \frac{\partial \eta}{\partial R_2} = \frac{(\tau + \frac{1}{\nu_2})R_2 + 1}{\beta_2 + \lambda_2} \qquad (15)$$

If $\beta_2, \lambda_2 \ll 1$ then we expect that the right hand side of (15) is always greater than unity, indicating that varying β_2 will have the greater impact on η. We give some examples for illustration.

We shall use the following typical parameter values; the time unit is one day.

τ = 19 days (the mean of 17 abd 21 days as suggested in section 4.3)

$\frac{1}{\nu_i}$ = 20 days for all i (Smalley et al., 1977)

$1/\lambda_2$ = 2 years - approximate period of time for immunity to wear off (section 1.5)

$1/\gamma_1$ = 7 days - mean time for antibodies to attain maximum level.

The results of the computation of η for the various values of the β_i and R_2 are given in Table 2. It is apparent from the last two columns that η increases more when the β_i are reduced than when R_2 is increased by a similar proportion (here by 10%). Also the percentage increases are greatest (least) when initially the β_i are large (small) and R_2 small (large). it would seem, therefore, that directing control measures at the β_i is a good overall strategy. Moreover, in practice, the resultant reduction in superfection would also mean an increase in R_2.

5.4 Simulation of Model

As mentioned at the beginning of the chapter, if we allow the infection rates to depend on time, then apparently the only way to proceed is by simulation. In this section we consider the special case in which the $\beta_i(t)$ are expressed as in equation (4) of Chapter 4. Using some judicious choices of rate parameter values we should be able to gain some insight of the progress of malaria transmission with the time of the year. First we give a brief description of the simulation procedures.

5.4.1 Method of Simulation

A FORTRAN program was written specifically to carry out the simulation. The procedure adopted closely follows the steps described in a book by Mihram (1972, page 232).

The simulation is performed with the aid of scheduling routines. These routines maintain an event list consisting of the next transition between states due to occur to each

TABLE 2

Values of η when the β_i and R_2 are separately changed by 10%

CASE		(day^{-1}) β_1	β_2	R_2	η	Increase Δ	%
$R_2 > \beta_1 > \beta_2$	(i)	0.0050	0.0025	0.0150	0.7570	-	-
	(ii)	0.0050	0.0025	0.0165	0.7633	0.0063	0.83
	(iii)	0.0045	0.0023	0.0150	0.7665	0.0095	1.25
$\beta_1 > R_2 > \beta_2$	(i)	0.0150	0.0075	0.0100	0.4676	-	-
	(ii)	0.0150	0.0075	0.0110	0.4843	0.0167	3.57
	(iii)	0.0135	0.0068	0.0100	0.4903	0.0227	4.85
$\beta_1 > \infty_2 > R_2$	(i)	0.0150	0.0100	0.0025	0.1791	-	-
	(ii)	0.0150	0.0100	0.0028	0.1947	0.0156	8.71
	(iii)	0.0135	0.0090	0.0025	0.2176	0.0385	21.49

In each case the values of β_1 and β_2 are both held fixed in (i) and (ii) and reduced by 10% in (iii). R_2 is held fixed in (i) and (iii), and increased by 10% in (ii).

Δ = difference between (i) and (ii) or (iii).

individual. This is necessary because of the non-Markovian nature of the process.

Initially, arbitrary numbers of individuals are assigned to each state, and the event list is filled with the random exit times from these states. Thus no attempt is made to start the process in equilibrium. Each cycle of the simulation then consists of :

(1) Advancing time to that of the earliest event on the list.

(2) Retrieving a code from the list describing the nature of the transition.

(3) Deleting the event from the list.

(4) Modifying the counts of individuals in each state according to the transition code.

(5) Generating the random sojourn time in the new state and the next state to be visited for the individual concerned.

(6) Scheduling the corresponding time and transition code on the event list.

Monitoring the progress of the simulation is achieved by scheduling a "report" event at regular intervals of time on the same event list.

The random number generators used are the uniform (GO5CAF), exponential (GO5[]BF) and gamma (GO5DGF) distribution generators of the NAG routine library (NAG, 1977). The scheduling and fetching subroutines, which implement the event list as a binary tree, were provided by Dr. Peter J. Green.

The points at which decisions about the next state have to be made in the model are shown in Figure 8. They are specifically of two types : getting infected or losing immunity level, and, retaining one's immunity level or gaining one in the infectious stage.

To generate non-negative random variables, given a sinusoidal hazard function, we use the cumulative distribution function transform method (Hammersley and Handscomb, 1964, Mihram, 1972). In this case the procedure reduces to solving (by Newton-Raphson method) the equation $\int_0^T \beta_i(u)du = a$ unit exponential random variable.

5.4.2 Results of Simulation

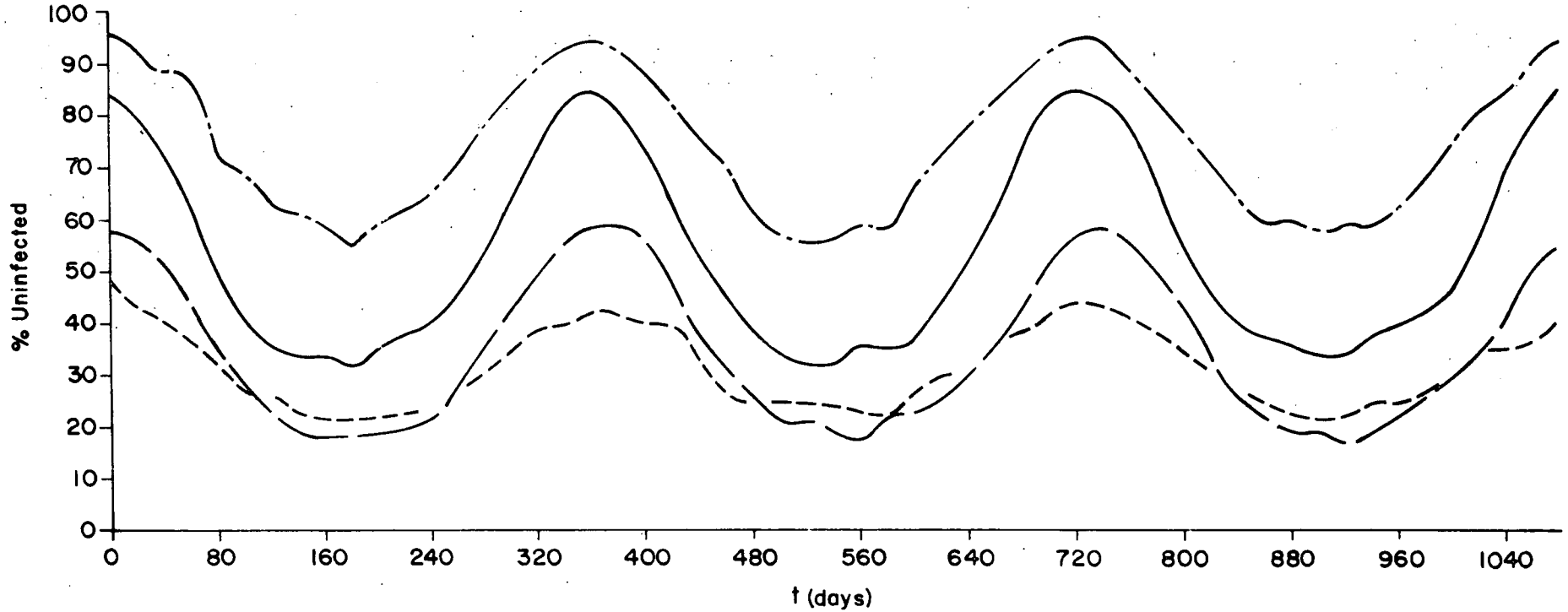
The time evolution of the prevalence of the disease has been simulated for the number of immunity classes k = 1 and 3. The results are presented graphically in figures 10 and 11.

The latent periods are again assumed constant and equal to 20 days for all immunity levels. The time distributions are as described in the last chapter.

Figure 10 shows the proportions of uninfected (or susceptibles) over several years for the case k = 1. The graph represents part of the simulation some years from the starting point in time. The simulation is of a population of 1000 initially assigned all to the susceptible phase. Whatever initial population distribution was used made little difference to the speed with which apparent equilibrium was attained.

The rate parameters were judiciously set initially

Fig. 10 THE EFFECT OF VARYING THE RATE PARAMETERS ON THE PROPORTION OF UNINFECTED PERSONS. k = 1



The four cases are based on the following parameter values:

	b	b'	α
———————	0·02	0·95	0·015
—— —— —	0·02	0·95	0·005
— — — — — —	0·02	0·60	0·005
—— - —— - —	0·01	0·95	0·025

as follows:

Infection rate : $b = 0.02$, $b' = 0.95$

Rate of loss of infectivity : $\nu = 0.1$

Recovery rate parameters : $\alpha = 0.015$, $r = 0.5$

The results of varying some of the parameter values are also shown in the figure.

The periodic forms of the variation in the susceptible population, and hence the infected, show very clearly in all cases in the figure. As expected, the oscillation for $b = 0.02$ is more pronounced than that for $b = 0.01$. In the latter case ($b = 0.01$) α was set at 0.025 to account for expected reduction in the effect of superinfection. From the figure we also note that the proportion of uninfected is consistently lower for higher values of b, a realistic response by the model. For instance, the maximum percentage of uninfected for $b = 0.02$ is about 85% while that for $b = 0.01$ is 95%. Their minima are about 33% and 55% respectively. In every case, the fall in the proportion of uninfected is slightly steeper than the rise. This is explained by the availability of a large pool of susceptibles accumulated over the dry season. After the peak of infection at the height of the wet season, relatively fewer uninfected individuals are left to be infected. This non-symmetric behaviour of the model is manifested later in Chapter 6 when rate parameters are estimated from actual field data, using simplified versions of the model.

Figure 11 is the result of the simulation for $k = 3$. Again we start with 1000 individuals, of which 200, 400 and

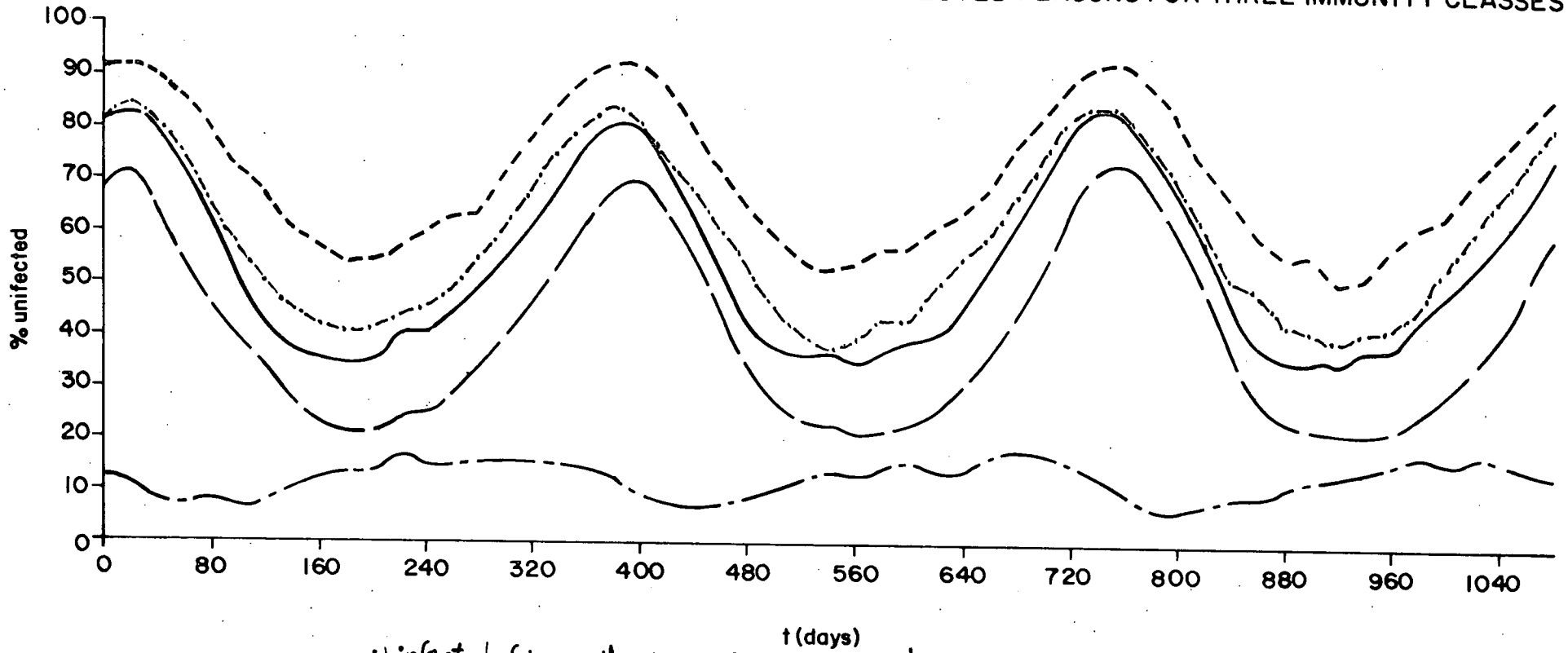
400 are assigned to immunity levels 1, 2 and 3, respectively, in the susceptible category.

The following table contains the parameter values used.

Immunity level		1	2	3
	rates			
Infection	b_i	0.02	0.01	0.01
	b'_i	0.95	0.75	0.5
Loss of infectivity	ν_i	0.1	0.15	0.15
Recovery	α_i	0.05	0.03	0.03
	r_i	0.5	0.5	0.5
Loss of immunity	λ_i	0.03	0.03	0.03
Gain of immunity	γ_i	0.15	0.03	0.03

The result shows that, taking the three immunity levels together, the proportion of the uninfected population varies periodically with time as in the case k = 1. Similar periodicity patterns are discernible when the uninfected immunity level 1 are considered separately. However, the number of uninfected in higher immunity levels actually increases during period of high infection and falls in the dry season.

Fig.II THE TIME DEPENDENCY OF THE PROPORTIONS OF UNINFECTED PERSONS FOR THREE IMMUNITY CLASSES



The fall in the percentage of all uninfected persons is slightly steeper than the rise. This is also the case for immunity class 1

One result of reducing the b_i and increasing α_i is shown in Figure 11. The proportion uninfected of the less immune (1) is now much higher than before. Although fewer people are infected, the general vulnerability of the population is now high. This observation is in agreement with reality as manifested by the Sri Lanka experience (Gilles, 1981). Reducing the rate of loss of immunity appears to check this vulnerability. The general level of infected population falls as it becomes more and more immune.

5.4.3 General Comments

Although the model proposed here may not yield usable results by analytic means, results from simulation of the model are in general agreement with some of what is known about malaria.

The sinusoidal hazard function for the distribution of time to infection seems to provide a picture of the process of the disease which agrees qualitatively with reality. One interesting result is that the rate of loss of immunity appears to be a more decisive factor than say reducing the infection rate. The efforts being expended to find a vaccine for malaria is thus underlined.

One major difficulty with the model lies in the lack of good data for validation. A better quantitative understanding is required of the various phases of the disease to provide more accurate estimates of rate parameters and allow the construction of more realistic models of transition time distributions.

CHAPTER SIX

SOME ANALYSIS OF THE GARKI DATA

6.1 Introduction

In the last chapter, we mentioned the lack of raw data to assess the adequacy of the semi-Markov model. In this chapter, we shall consider simplified versions of the model to analyse some of the data from the Garki project described in Chapter 1 (section 1.6). In particular, we shall be concerned with the estimation of 'apparent' incidence and 'apparent' recovery rates of malaria and other related quantities, within different age groups. These rates are said to be 'apparent' because the distinction, for instance, between new infections and relapses from old infections is not made in the Garki data. Sometimes incidence rates may be referred to as infection rates.

We begin by introducing a general form of the transition probabilities for a finite state Markov process. We then focus on the special case of a 2-state Markov process, which is essentially that of an alternating renewal process (section 2). We also derive expressions for equilibrium probabilities for a time-inhomogeneous version of such a process when the elements of the intensity matrix vary periodically (subsection 2.2). The concept of embedding processes in a class of continuous-time Markov chain is discussed in subsection 2.3. In section 3, we estimate the incidence and recovery rates by assuming exponential (subsection 3.1) and some other distributions (subsections 3.2 and 3.3) for the times to recovery, followed by a discussion of the results

(subsection 3.4). Inhomogeneous Markov chain models are considered in section 4, and applied to two sets of data. The results are discussed and compared with those from models of the previous sections in the concluding section (5).

The analysis in subsection 3.1 is similar to that in Bekessy et al. (1976) and Singer and Cohen (1980). However, it was conceived and largely carried out without prior knowledge of these papers. We have, nonetheless, incorporated a number of ideas from them. For example, we had started by using age group range of 5 years for the subjects in the Garki project. But the ages that were recorded at Garki, especially among the older subjects, were usually estimates with a preference for multiples of 5 years. The grouping advocated by Bekessy et al., and which we now use, attempts to minimize this bias (Molineaux et al., 1980, pages 232-233). We have also used their set of data (Table 5) for most of our analysis. The condition of embedding a stochastic matrix, generated by a given set of data, in a class of continuous-time Markov chain is that from the paper of Singer et al. (1980). Other ideas from these two papers that have been used in connection with subsection 3.1, are appropriately referenced.

6.2 Transition Probabilities of Markov Processes

Following Singer et al., we shall assume that a two parameter family of finite NxN stochastic matrices, which we denote by $\{ \underline{P}(s,t) \}$, $o \leq s \leq t < t_o < \infty$, satisfies the following conditions (Goodman, 1970);

(a) each element, $p_{ij}(s,t)$ of $\underline{P}(s,t)$ is continuous in (s,t), $i,j = 1,2,\ldots,N$;

(b) $\underline{P}(s,t) = \underline{I}$, the identity matrix, if and only if $s = t$;

(c) the Chapman-Kolomogonov equations hold; namely

$$\underline{P}(s,t) = \underline{P}(s,v)\ \underline{P}(v,t) \qquad o \leq s \leq v \leq t$$

The matrix $\underline{P}(s,t)$ also satisfies the Kolomogorov forward and backward equations, written in matrix notation as

$$\begin{aligned} &\text{(i)} \quad \frac{\partial}{\partial t}\ \underline{P}(s,t) = \underline{P}(s,t)\ \underline{A}(t), \quad \underline{P}(t,t) = \underline{I}; \\ &\text{(ii)} \quad \frac{\partial}{\partial s}\ \underline{P}(s,t) = -\underline{A}(s)\ \underline{P}(s,t), \quad \underline{P}(s,s) = I \end{aligned} \tag{1}$$

provided $\frac{\partial}{\partial s}\ \underline{P}(s,t)$ and $\frac{\partial}{\partial t}\ \underline{P}(s,t)$ exist for all $s \leq t$. The elements $a_{ij}(u)$ of the NxN matrix $\underline{A}(u)$ are bounded measurable functions such that, for every $u \geq o$,

$$\begin{aligned} &\text{(i)} \quad a_{ii}(u) \leq o\ , \quad i = 1,2,\ldots,N \\ &\text{(ii)} \quad a_{ij}(u) \geq o\ , \quad i \neq j = 1,2,\ldots,N \\ &\text{(iii)} \quad \sum_{j=1}^{N} a_{ij}(u) = o \quad i = 1,2,\ldots,N \end{aligned} \tag{2}$$

The matrix $\underline{P}(s,t)$ may be viewed as describing an inhomogeneous continuous-time Markov chain, with the elements $p_{ij}(s,t)$ as the transition probabilities from state i at time s to state j at time t. $\underline{A}(u)$ is then the instantaneous intensity matrix of the process.

6.2.1 The 2-state Markov process

In the case N = 2, $\underline{P}(s,t)$ may be written as,

$$\underline{P}(s,t) = \begin{bmatrix} 1 - p_{12}(s,t) & p_{12}(s,t) \\ p_{21}(s,t) & 1-p_{21}(s,t) \end{bmatrix}, \quad o \leq s \leq t \qquad (3)$$

where $o \leq p_{12}(s,t), p_{21}(s,t) \leq 1$, and the matrix $\underline{A}(\cdot)$ may be expressed in the form

$$\underline{A}(t) = \begin{bmatrix} - \beta(t) & \beta(t) \\ \alpha(t) & -\alpha(t) \end{bmatrix} \qquad (4)$$

where $\alpha(t), \beta(t) > o$.

Goodman (1970) has shown that if

$$\sup_i \int_s^t |a_{ii}(v)| \, dv < \infty \quad , \text{ for all s and t, } s < t, \qquad (5)$$

in addition to the conditions in (2), then the system of equations (1(i)) possesses a general solution $\underline{P}(s,t)$ that has the properties (a), (b) and (c). The solution is also absolutely continuous in s and in t, and satisfies the system of backward equations (1(ii)).

In the case N = 2, suppose $\alpha(\cdot)$ and $\beta(\cdot)$ satisfy (5). Then using (3) and (4) we may write the forward equations in the form

$$\text{(i)} \quad \frac{\partial}{\partial t} p_{12}(s,t) = \beta(t)(1 - p_{12}(s,t)) - \alpha(t)p_{12}(s,t) \; ;$$

$$\text{(ii)} \quad \frac{\partial}{\partial t} p_{21}(s,t) = \alpha(t)(1 - p_{21}(s,t)) - \beta(t)p_{21}(s,t) \qquad (6)$$

Solving the above partial differential equations, we obtain

$$\text{(i)}\quad p_{12}(s,t) = \int_s^t \beta(v)\exp\left(-\int_v^t (\alpha(u)+\beta(u))du\right) dv$$

$$\text{(ii)}\quad p_{21}(s,t) = \int_s^t \alpha(v)\exp\left(-\int_v^t (\alpha(u)+\beta(u))\,du\right)dv \qquad (7)$$

The inequalities

$$p_{11}(s,t)\, p_{22}(s,t) \geq \det \underline{P}(s,t) > o, \qquad (8)$$

where det $\underline{P}(s,t)$ denotes the determinant of $\underline{P}(s,t)$, follow trivially from

$$\begin{aligned}\det \underline{P}(s,t) &= 1 - p_{12}(s,t) - p_{21}(s,t)\\ &= \exp\left(-\int_s^t (\alpha(u)+\beta(u))\,du\right)\\ &> o,\end{aligned}$$

and

$$\begin{aligned}p_{11}(s,t)\, p_{22}(s,t) &= (1 - p_{12}(s,t))\,(1 - p_{21}(s,t))\\ &= \det \underline{P}(s,t) + p_{12}(s,t)\, p_{21}(s,t)\\ &\geq \det \underline{P}(s,t).\end{aligned}$$

Clearly the equality holds only if either state is absorbing.

Det $\underline{P}(s,t)$ is interpreted (Karlin and McGregor, 1959) as the probability that two individuals, each starting at time s in states 1 (susceptible or negative state) and 2 (infected or positive state) respectively, and evolving according to the law of the process, continue to occupy their original states throughout the interval (s,t). The left side inequality in (8) necessarily holds because the product, $p_{11}(s,t)\, p_{22}(s,t)$, includes the probability of an even number

of transitions during the time interval (s,t). We note that det $\underline{P}(s,t) \to o$ as $|s - t| \to \infty$, which means the tendency to remain in one's initial state decreases as the time of continuous occupation of the state increases.

6.2.2 Periodic intensity matrices and equilibrium solutions

It may be of interest to suppose that the elements of the intensity matrix $\underline{A}(t)$ vary periodically with period y. In that case, the equilibrium probabilities of the process, if they exist, depend on t.

Let $\underline{\Pi}(t) = (\Pi_1(t), \Pi_2(t))$ be the row vector of the equilibrium probabilities. Then the equilibrium equations are of the form

$$\underline{\Pi}(s)\,\underline{P}(s,t) = \underline{\Pi}(t) \tag{9}$$

with $\Pi_1(s) + \Pi_2(s) = 1$

Or

$$\begin{aligned} &\text{(i)} \quad \Pi_1(s)(1-p_{12}(s,t)) + \Pi_2(s)p_{21}(s,t) = \Pi_1(t) \\ &\text{(ii)} \quad \Pi_1(s)p_{12}(s,t) + \Pi_2(1-p_{21}(s,t)) = \Pi_2(t), \quad s < t \end{aligned} \tag{10}$$

Let $t = s + ky$ $k = o,1,2,\ldots.$

We expect that $\Pi_i(s + ky) = \Pi_i(s)$ (because of periodicity of $\underline{A}(\cdot)$) for all i, then (10) may be written as

$$\begin{aligned} &\text{(i)} \quad \Pi_1(s)(1 - p_{12}(s,s + ky)) + \Pi_2(s)p_{21}(s,s + ky) = \Pi_1(s) \\ &\text{(ii)} \quad \Pi_1(s)p_{12}(s,s + ky) + \Pi_2(s)(1 - p_{21}(s,s + ky) = \Pi_2(s) \end{aligned} \tag{11}$$

Solving for $\Pi_1(s)$ and $\Pi_2(s)$, we obtain

$$\text{(i)} \quad \Pi_1(s) = \frac{p_{21}(s,s+ky)}{p_{12}(s,s+ky)+p_{21}(s,s+ky)}$$

$$\text{(ii)} \quad \Pi_2(s) = \frac{p_{12}(s,s+ky)}{p_{12}(s,s+ky)+p_{21}(s,s+ky)} \tag{12}$$

for all integer k.

The denominator of (12) may be obtained from (7) to be of the form

$$1 - \frac{I(s+ky)}{I(s)} \quad ,$$

where $I(s) = \exp(-\int_o^s (\alpha(v)+\beta(v))\, dv)$. And since $\alpha(v)$ and $\beta(v)$ are both periodic, we have

$$I(s+ky) = \exp(-\int_o^{s+ky} (\alpha(v)+\beta(v))\, dv)$$

$$= I(s)I(y)^k \tag{13}$$

so that the denominator of (12) becomes $1 - I(y)^k$. We note that, since $\alpha(v)$, $\beta(u) > o$, it follows trivially from (13) that, if we set $s = o$ and $k = 1$, we have $o < I(y) < 1$.

We also have

$$p_{21}(s,s+ky) = \int_s^{s+ky} \alpha(v)\exp(-\int_v^{s+ky} (\alpha(u)+\beta(u))du)\, dv$$

$$= \sum_{i=1}^{k} \int_{s+(i-1)y}^{s+iy} \alpha(v)\exp(-\int_v^{s+iy} (\alpha(u)+\beta(u))du)dv$$

$$= (1 + I(y) + I(y)^2 + \ldots\ldots + I(y)^{k-1})$$

$$x \int_s^{s+y} \alpha(v)\exp(-\int_v^{s+v} (\alpha(u)+\beta(u))du)\, dv$$

$$= \frac{(1-I(y)^k)}{1-I(y)}\ p_{21}(s,s+y) \tag{14}$$

Similarly we may obtain the result

$$p_{12}(s,s+ky) = \frac{(1-I(y)^k)}{1-I(y)}\, p_{12}(s,s+y) \qquad (15)$$

By substituting $1-I(y)^k$ for the denominator and (14) for the numerator of the right side of (12(i)), we obtain

$$(i) \qquad \Pi_1(s) = \frac{p_{21}(s,s+y)}{1-I(y)}$$

and from $\Pi_2(s) = 1 - \Pi_1(s)$, (16)

$$(ii) \qquad \Pi_2(s) = \frac{p_{12}(s,s+y)}{1-I(y)}$$

These results may also be obtained as limiting probabilities, directly from (14 and (15) as $k \to \infty$ ($o < I(y) < 1$).

6.2.3 The Embeddability Condition

According to Singer and Cohen (1980), a finite stochastic matrix $\underline{P}$ is said to be embeddable in a continuous-time Markov chain if there exists a two-parameter family of stochastic matrices $\underline{P}(s,t)$ satisfying conditions (a), (b) and (c), given at the beginning of this section, in addition to $\underline{P}(o,1) = \underline{P}$. A necessary condition for embeddability of matrix $\underline{P}$ of any order has been proved by Goodman (1970) to be $\det\underline{P} > o$. This condition has been shown to be also sufficient for 2 x 2 stochastic matrices (homogeneous case), and in this case is equivalent to trace $\underline{P} > 1$ (Singer et al. 1980).

As pointed out by Singer et al., such a matrix $\underline{P}$ has the disadvantage of being embeddable in uncountably many inhomogeneous continuous-time Markov chains. But this same

aspect may be used to advantage since it allows for some degree of flexibility in the choices of suitable algebraic forms for $\alpha(\cdot)$ and $\beta(\cdot)$, based either on some nice mathematical properties or some known characteristics of the process under consideration. The matrix is, however, embeddable in a unique homogeneous chain (ibid). It must also be pointed out that non-Markovian processes are capable of generating stochastic matrices $\underline{P}$ such that det $\underline{P} > o$.

They proceed to develop hypotheses to test, a priori, whether or not a set of data may be embedded in a homogeneous continuous time Markov chain. The null hypothesis, H_o: det $\underline{P} > o$, is tested against the alternative hypothesis, H_1 : det $\underline{P} < o$. A three-way decision rule is used, based on some suitably derived parameters $\delta_i > o$, $i = 1,2$.

Thus, let $\hat{\underline{P}}$ be an estimate of $\underline{P}$. Then, if

(i) det $\hat{\underline{P}} > \delta_1$, accept H_o ;

(ii) det $\hat{\underline{P}} < -\delta_2$, accept H_1

(iii) $-\delta_2 \leq$ det $\hat{\underline{P}} \leq \delta_1$, there is inadequate evidence either to accept or reject H_o, probably an indication of insufficient data. They do not say, however, what should be done in such a case. The numbers δ_1 and δ_1 are derived from asymptotic sampling distribution of $\underline{P}$ for given probabilities of rejection, α_1 and α_2 respectively.

6.3 Estimating the Incidence and Recovery Rates

The 2-state model outlined so far is essentially a simplified version of the semi-Markov model described in Chapter 4. We have reduced the number of states from four to two within each immunity category by combining the latent, infectious and non-infectious states to form the positive or infected state. The susceptible state is also known as the negative state. A person is in state 1 or the negative state if he has no microscopically detectable parasitamia in his blood, otherwise he is in state 2, or the positive state. We note that state 1 is not a susceptible state in the real sense, because the mere absence of parasitamia (checked by microscope) is no proof of being free from the disease (Cohen, 1979). In addition to combining some of the states, we have assumed there are no transitions between immunity levels. These levels now correspond to age-groups. Our basic assumption is still that for each individual, the underlying stochastic process between two successive observations is a 2-state semi-Markov process, or alternating renewal process.

6.3.1 Homogeneous Markov Chains

The simplest version of the 2-state model, herein referred to as model A, is one where the times of transition between states are assumed to be exponentially distributed. The model thus describes a 2-state homogeneous continuous-time Markov process, where $\underline{P}(s,t)$ depends only on the length, t-s, of the time interval (s,t) and so is equivalent to $\underline{\Phi}(t-s)$, $o \leq s \leq t$, (section 3.3).

Because of the homogeneity assumption, the elements of the intensity matrix $\underline{A}(\cdot)$, are constants so that equations (7) now become

$$\text{(i)} \quad \phi_{12}(t-s) = \frac{\beta}{\alpha+\beta}\{\ 1 - \exp(-(\alpha+\beta)(t-s))\ \}$$

$$\text{(ii)} \quad \phi_{21}(t-s) = \frac{\alpha}{\alpha+\beta}\{\ 1 - \exp(-(\alpha+\beta)(t-s))\ \} \qquad (17)$$

As mentioned earlier (section 1), model A is identical to that applied by Bekessy et al to baseline data of the Garki project. This set of data is reproduced in Table 3. The data are transition frequencies, n_{ij} of persons observed in state i at one survey and in state j at the next survey, i,j = 1,2. Each entry is a 2 x 2 array

$$\begin{bmatrix} n_{11} & n_{12} \\ n_{21} & n_{22} \end{bmatrix},$$

classified by age group and the pair of successive surveys. The age groups, seven in number, used here and in all subsequent discussions are : less than one year (o < 1), 1 - 4, 5 - 8, 9 - 18, 19 - 28, 29 - 43 and 44 years and over. Five pairs of surveys are used, which are (3,4), (4,5), (5,6), (6,7) and (7,8); they are respectively 68, 78, 81, 76 and 70 days apart. The ages are all measured from survey 1.

The 2 x 2 tables are all treated as independent samples with separate parameter estimates. To carry out the estimation procedure, we assume the two rows of each entry to be independent binomial observations. The logarithm of

TABLE 3

SURVEY PAIRS	AGE GROUPS											
	0 < 1		1 - 4		5 - 8		9 - 18		19 - 28		29 - 43		44+	
(3,4)	61	15	38	21	27	47	111	70	378	100	810	201	509	113
Dry season	12	42	46	448	58	484	87	270	86	105	163	107	99	61
(4,5)	21	66	20	71	20	76	102	135	341	200	706	393	401	239
Wet season	6	24	31	459	43	568	65	375	98	152	174	179	98	86
(5,6)	8	31	14	36	18	52	79	90	273	160	635	277	354	140
Wet season	6	56	33	451	68	581	122	415	157	199	324	245	190	121
(6,7)	28	11	16	29	22	49	86	61	280	65	716	126	433	71
Dry season	12	67	32	422	66	484	106	274	174	126	314	148	173	70
(7,8)	63	6	21	23	26	54	120	59	355	81	845	169	528	108
Dry season	17	54	21	386	46	488	81	241	87	92	161	101	91	45

Malaria Parasitological Transition Data : The entries in the 2 x 2 arrays are $\begin{matrix} n_{11} & n_{12} \\ n_{21} & n_{22} \end{matrix}$, where n_{ij} are the observed number of persons in state i at one survey and in state j at the next survey. Source : Bekessy et al. (1976).

the likelihood function of an entry for the ℓth survey pair is given by

$$L(\alpha, \beta; t_\ell) = \sum_{i,j}^{2} n_{ij} \log_e \phi_{ij}(t_\ell) + \sum_{i=1}^{2} \log_e \binom{n_{i\bullet}}{n_{ii}} \qquad (18)$$

where t_ℓ is the number of days between the ℓth survey pair, $\ell = 1,2\ldots,5$, $n_{i\bullet} = n_{i1} + n_{i2}$ and

$$\binom{n_{i\bullet}}{n_{ii}} = \frac{n_{i\bullet}!}{n_{ii}!\,(n_{i\bullet}-n_{ii})!}$$

The maximum likelihood estimate of $\phi_{ij}(t_\ell)$ is given by $\hat{\phi}_{ij}(t_\ell) = n_{ij}/n_{i\bullet}$, $i,j = 1,2$. We obtain the maximum likelihood estimates of α and β of each entry by solving (17), using the estimates of $\phi_{12}(t_\ell)$ and $\phi_{21}(t_\ell)$, with t-s replaced by t_ℓ.

$$\text{(i)} \quad \hat{\alpha} = -\frac{\hat{\phi}_{21}\log_e(1-\hat{\phi}_{12}-\hat{\phi}_{21})}{t_\ell(\hat{\phi}_{12}+\hat{\phi}_{21})}$$

$$\text{(ii)} \quad \hat{\beta} = -\frac{\hat{\phi}_{12}\log_e(1-\hat{\phi}_{12}-\hat{\phi}_{21})}{t_\ell(\hat{\phi}_{12}+\hat{\phi}_{21})} \qquad (19)$$

where the arguments of $\phi_{ij}(\cdot)$ are suppressed.

The argument of the logarithmic term in (19) is $\det \hat{\underline{\Phi}}$, and must be positive for α and β to be estimated. This is the embeddability condition pointed out in sub-section 2.3. Bekessy et al. have suggested that a negative value of $\det \hat{\underline{\Phi}}$ may mean the data are generated by an inhomogeneous process which may still be Markovian. However, by the embeddability criterion this is not necessarily the case, a point

noted by Singer and Cohen (1980).

Provided det $\underline{\hat{\phi}} > 0$, the estimated covariance matrix of $\hat{\alpha}$ and $\hat{\beta}$, denoted by $C(\hat{\alpha},\hat{\beta})$, is given by the inverse of the Fisher information matrix, computed at $\alpha = \hat{\alpha}$ and $\beta = \hat{\beta}$.

$$C(\hat{\alpha},\hat{\beta}) = - \begin{bmatrix} \frac{\partial^2 L}{\partial \alpha^2} & \frac{\partial^2 L}{\partial \alpha \partial \beta} \\ \frac{\partial^2 L}{\partial \alpha \partial \beta} & \frac{\partial^2 L}{\partial \beta^2} \end{bmatrix}^{-1}_{\alpha=\hat{\alpha} \quad \beta=\hat{\beta}} \tag{20}$$

The elements of the information matrix are:

$$\text{(i)} \quad -\frac{\partial^2 L}{\partial \alpha^2} = \{S_1(\hat{\phi}_{12}V + \hat{\phi}_{21}U)^2 + S_2(\hat{\phi}_{12}V - \hat{\phi}_{12}U)^2\} / (S_1 S_2 U^2 V^2)$$

$$\text{(ii)} \quad -\frac{\partial^2 L}{\partial \beta^2} = \{S_2(\hat{\phi}_{21}V + \hat{\phi}_{12}U)^2 + S_1(\hat{\phi}_{21}V - \hat{\phi}_{21}U)^2\} / (S_1 S_2 U^2 V^2) \tag{21}$$

$$\text{(iii)} \quad -\frac{\partial^2 L}{\partial \alpha \partial \beta} = (U - V)\{S_2 \hat{\phi}_{12}(\hat{\phi}_{21}V + \hat{\phi}_{12}U) + S_1 \hat{\phi}_{21}(\hat{\phi}_{12}V + \hat{\phi}_{21}U)\} / (S_1 S_2 U^2 V^2) ,$$

where

$$\text{(i)} \quad U = -\log_e(1 - \hat{\phi}_{12} - \hat{\phi}_{21})/t_\ell$$

$$\text{(ii)} \quad V = (\hat{\phi}_{12} + \hat{\phi}_{21}) / \{(1 - \hat{\phi}_{12} - \hat{\phi}_{21})t_\ell\} \tag{22}$$

$$\text{(iii)} \quad S_i = \hat{\phi}_{ij}(1 - \hat{\phi}_{ij})/n_{i\bullet} , \quad i \neq j = 1,2$$

Using equations (21), the explicit formula for the elements of the covariance matrix (20) are then,

$$\text{(i)} \quad \text{var}(\alpha) = \{S_2(\hat{\phi}_{21}V + \hat{\phi}_{12}U)^2 + S_1(\hat{\phi}_{12}V - \hat{\phi}_{21}U)^2\} / (\hat{\phi}_{12} + \hat{\phi}_{21})^4$$

$$\text{(ii)}\quad \text{var}(\hat{\beta}) = \{S_1(\hat{\phi}_{12}V+\hat{\phi}_{21}U)^2+S_2(\hat{\phi}_{21}V-\hat{\phi}_{12}U)^2\}\ /(\hat{\phi}_{12}+\hat{\phi}_{21})^4 \qquad (23)$$

$$\text{(iii)}\quad \text{cov}(\hat{\alpha},\hat{\beta}) = (V-U)\{S_1\hat{\phi}_{21}(\hat{\phi}_{12}V+\hat{\phi}_{21}U)+S_2\hat{\phi}_{12}(\hat{\phi}_{21}V+\hat{\phi}_{12}U)\}\ /(\hat{\phi}_{12}+\hat{\phi}_{21})^4$$

Equations (23) have also been derived by Bekessy et al. ((i) and (ii) only) and Mueller (in Singer and Cohen, 1980).

The results of applying model A to the data in Table 3 are listed in Table 4. These are the estimates of α and β (columns 3 and 4 respectively) and their respective estimated standard errors (columns 3 and 4 (in brackets)) obtained from the estimated covariance matrix. The table also contains the corresponding estimates for det$\underline{\Phi}$ (column 6), which may be thought of as some measure of embeddability. Clearly, none of the entries of Table 5 violates the embeddability condition, although about four values of det $\underline{\Phi}$ are relatively close to zero (survey pairs (4,5) and (5,6) for both age groups of less than one year ($o < 1$) and 44+ years). However, before we proceed to discuss these results we would like to consider other models that may also provide good, if not better, descriptions of the data.

6.3.2 Other non-Markovian Models

It may be argued that, while it is fairly reasonable to assume that incidence times are exponentially distributed, the duration of an infection is not entirely memoriless. Hence, in this section we explore some cases in which time to recovery is assumed to have distributions other than exponential. In particular, we consider the case where recovery time has a gamma distribution with parameters $r > 1$, and α, as suggested earlier in Chapter 4. More specifically, we

TABLE 4

Malaria Recovery and Incidence rates and their standard errors estimated by model A, plus det $\underline{\Phi}$.

Survey pairs	Age groups (years)	Recovery rate ($10^{-4}day^{-1}$)	(s.e.)	Incidence rate ($10^{-4}day^{-1}$)	(s.e.)	% correl. coeff.	det $\hat{\underline{\Phi}}$
(3,4)	0 <1	48	(13)	38	(10)	28	0.58
	1-4	18	(3)	70	(16)	4	0.55
	5-8	29	(5)	171	(30)	60	0.26
	9-18	57	(7)	90	(12)	50	0.37
	19-28	108	(14)	50	(6)	53	0.34
	29-43	179	(19)	59	(6)	72	0.20
	44+	183	(25)	54	(7)	7	0.20
(4,5)	0<1	85	(76)	323	(194)	95	0.04
	1-4	18	(4)	220	(34)	60	0.16
	5-8	21	(4)	233	(37)	64	0.14
	9-18	33	(5)	129	(13)	53	0.28
	19-28	95	(13)	89	(9)	69	0.24
	29-43	141	(17)	102	(10)	97	0.15
	44+	178	(37)	125	(21)	91	0.09
(5,6)	0<1	30	(16)	245	(73)	69	0.11
	1-4	17	(4)	175	(35)	58	0.21
	5-8	29	(5)	204	(39)	79	0.15
	9-18	53	(7)	123	(17)	66	0.24
	19-28	112	(13)	94	(11)	74	0.19
	29-43	166	(17)	89	(9)	82	0.13
	44+	190	(28)	88	(14)	85	0.11
(6.7)	0<1	26	(8)	49	(15)	28	0.57
	1-4	16	(4)	149	(31)	54	0.29
	5-8	32	(6)	186	(35)	71	0.19
	9-18	63	(8)	93	(14)	60	0.31
	19-28	145	(15)	47	(7)	62	0.23
	29-43	191	(16)	42	(5)	65	0.17
	44+	211	(24)	42	(7)	70	0.15
(7,8)	0<1	41	(10)	15	(6)	19	0.67
	1-4	11	(3)	111	(25)	37	0.43
	5-8	23	(4)	181	(30)	56	0.24
	9-18	54	(7)	71	(10)	44	0.42
	19-28	115	(15)	44	(6)	50	0.33
	29-43	171	(18)	46	(5)	66	0.22
	44+	208	(32)	53	(8)	75	0.16

discuss the cases r = 2,3, assuming the process is in equilibrium.

Essentially, these models are examples of equilibrium alternating renewal processes. Regardless of the distribution forms, if we define $\Pi_{ij}(\cdot)$ as

$$\Pi_{ij}(t) = \lim_{s\to\infty} \text{Prob}\{ Z_{s+t} = j \mid Z_s = i\}, \quad i,j = 1,2 \tag{24}$$

then Cox (1970, pages 85-86) has shown that

$$\Pi_{ij}(t) = \eta_j - \frac{\omega(t)}{\mu_i} \qquad i \neq j = 1,2 \tag{25}$$

where, as before, η_j is the limiting probability of being in state j and the function $\omega(t)$ has Laplace transform $\omega^*(\theta)$ given by

$$\omega^*(\theta) = \frac{\eta_2 \mu_1}{\theta} - \frac{(1-F_{12}^{**}(\theta))(1-F_{21}^{**}(\theta))}{\theta^2(1-F_{12}^{**}(\theta)\, F_{21}^{**}(\theta))} \tag{26}$$

We note that here the transition probabilities of the embedded Markov chain are simply $p_{11} = p_{22} = o$ and $p_{12} = p_{21} = 1$.

In our particular case, we have

$$F_{12}^{**}(\theta) = \frac{\beta}{\beta+\theta}, \quad F_{21}^{**}(\theta) = \left(\frac{\alpha}{\alpha+\theta}\right)^r \tag{27}$$

with $\mu_1 = \frac{1}{\beta}$ and $\mu_2 = \frac{r}{\alpha}$. We first consider the case r = 2, which we identify as model B. This is followed by model C, where r = 3.

(a) Model B (r = 2)

We substitute (27) with r = 2 into equation (26) and invert the resulting Laplace transform to obtain $\omega(t)$. The result is substituted into (25), and after some

simplification and rearrangement,

$$\Pi_{ij}(t) = \eta_j(1 - g(\alpha,\beta,2;t)) \qquad i \neq j = 1,2 \tag{28}$$

where

$$g(\alpha,\beta,2;t) = \begin{cases} \exp(u_1\alpha t)\{2u_2\cos(u_2\alpha t)+(3+2u_1)\sin(u_2\alpha t)\}/2u_2 \\ \qquad\qquad o < \rho < 8 \\ \exp(u_1\alpha t)\{(3+2(u_1+u_2))\exp(u_2\alpha t) \\ \qquad +(3+2(u_1-u_2))\exp(-u_2\alpha t)\}/u_2 \\ \qquad\qquad \rho \geq 8 \end{cases} \tag{29}$$

with

(i) $u_1 = -(4+\rho)/4$

$$\text{(ii)}\quad u_2 = \begin{cases} (8\rho - \rho^2)/4 & o < \rho < 8 \\ (\rho^2 - 8\rho)/4 & \rho \geq 8 \end{cases} \tag{30}$$

and ρ is the ratio of the mean recovery and infection rates.

If we apply model B to each of the 2 x 2 entries of Table 3, we obtain, by the maximum likelihood method, the following relations

$$\hat{\Pi}_{ij} = \hat{\eta}_j(1 - \hat{g}) = n_{ij}/n_{i\bullet} \quad i \neq j = 1,2 \tag{31}$$

where all arguments are suppressed. Equations (31) immediately lead to

$$\text{(i)}\quad \hat{\eta}_i = \frac{n_{ji}n_{i\bullet}}{n_{12}n_{2\bullet} + n_{21}n_{1\bullet}} \qquad i \neq j = 1,2$$

$$\text{(ii)}\quad \hat{g} = 1 - n_{12}/n_{1\bullet} - n_{21}/n_{2\bullet}, \tag{32}$$

from which we also obtain

$$\hat{\rho} = (n_{12} n_{2\bullet})/(n_{21} n_{1\bullet}) \qquad (33)$$

From the expression of $g(\alpha,\beta, 2;t)$ it is clear that estimates for α and β can only be obtained by numerical methods. To this end, it is required only to solve equation (32(ii)) for α, using for instance, the Newton-Raphson method. Estimate for β is then obtained from the relation $2\beta = \rho\alpha$. The results of the computation are listed in Table 5. The asymptotic covariance matrix is obtained by evaluating the derivatives by finite difference, and we use Romberg's interpolation method (Sheid[c], 1968, page 108; Fox and Meyers, 1968, page 176), to reduce truncation and rounding off errors.

A quick comparison of the results in Tables 4 and 5 shows that the estimates of the rates of model B are consistently lower than those of model A. But before a detail comparison, let us consider the case r = 3.

(b) Model C (r = 3)

In this model, the substitution of (27) into equations (26) for r = 3, and the result of inverting the Laplace transform, yield expressions identical to equation (28). However, we now have $g(\alpha,\beta, 3;t)$ expressed as

$$g(\alpha,\beta, 3;t) = \sum_{i=1}^{3} \frac{3a_i^2 + 8a_i + 6}{3(a_{i+1}-a_i)(a_{i+2}-a_i)} \exp(a_i \alpha t) \qquad (34)$$

where addition of subscripts is modulo 3 if it is greater than 3. The αa_i, i = 1,2,3, are the cube roots of the Laplace

TABLE 5

Malaria Recovery and Incidence rates and their standard errors (s.e.) estimated by model B.

Survey pairs	Age groups (years)	Recovery rate (10^{-4} day^{-1}) $\hat{\alpha}/2$	(s.e.)	Incidence rate (10^{-4} day^{-1}) $\hat{\beta}$	(s.e.)	% correl. coeff
(3,4)	0<1	38	(10)	34	(9)	17
	1-4	17	(2)	66	(10)	31
	5-8	26	(3)	153	(23)	61
	9-18	48	(6)	77	(9)	34
	19-28	87	(9)	41	(5)	37
	29-43	135	(11)	45	(4)	58
	44+	138	(15)	41	(5)	57
(4,5)	0<1	59	(13)	224	(42)	44
	1-4	16	(2)	198	(40)	96
	5-8	18	(2)	207	(37)	96
	9-18	29	(4)	113	(14)	44
	19-28	75	(7)	71	(7)	53
	29-43	105	(8)	76	(6)	69
	44+	128	(17)	89	(12)	82
(5,6)	0<1	25	(6)	206	(48)	61
	1-4	15	(2)	160	(33)	94
	5-8	25	(3)	175	(26)	76
	9-18	44	(5)	103	(11)	50
	19-28	86	(8)	72	(7)	61
	29-43	121	(10)	65	(6)	74
	44+	135	(18)	63	(9)	79
(6,7)	0<1	24	(6)	45	(13)	20
	1-4	15	(2)	138	(27)	84
	5-8	28	(4)	161	(23)	59
	9-18	52	(6)	78	(9)	42
	19-28	111	(13)	36	(5)	52
	29-43	141	(14)	31	(3)	61
	44+	154	(22)	30	(5)	68
(7,8)	0<1	37	(11)	13	(3)	10
	1-4	10	(1)	106	(24)	78
	5-8	21	(2)	164	(27)	68
	9-18	46	(5)	58	(7)	28
	19-28	92	(10)	35	(4)	38
	29-43	130	(11)	35	(3)	53
	44+	153	(19)	39	(5)	65

transform $\omega^*(\theta)$, and the a_i are given by

$$\text{(i)} \quad a_1 = c\text{-}d/c - 1 - \rho/9$$

$$\text{(ii)} \quad a_2 = wc - w^2 d/c - 1 - \rho/9 \qquad (35)$$

$$\text{(iii)} \quad a_3 = w^2 c - w\, d/c - 1 - \rho/9$$

where

$$\text{(i)} \quad c = \{ \frac{\rho}{18} (K - L/81) \}$$

$$\text{(ii)} \quad d = \frac{\rho}{81} (9 - \rho)$$

$$\text{(iii)} \quad w = \frac{1}{2} + \frac{i\sqrt{3}}{2}$$

$$\text{(iv)} \quad K = 1 + \frac{\rho(\rho-14)}{81}$$

$$\text{(v)} \quad L = 2\rho^2 - 27\rho + 243$$

We still have the same expression for the estimates of $g(\alpha, \beta, 3;t)$ and ρ in terms of n_{ij} given by (32(ii)) and (33) respectively.

Again we must resort to the Newton-Raphson technique to calculate estimates of α and β. As in model B, we first solve for α and then obtain β from the relation $3\beta = \rho\alpha$. The estimates of the parameters from the data in Table 3 appear in Table 6. Attempts to obtain their standard errors by computing the asymptotic covariance matrices numerically, as in model B, produced rather inconsistent results. This is probably because of truncation and rounding off errors which even the use of such refinement as Romberg's interpolation technique could not reduce significantly.

One way to get round this difficulty is to estimate the variances and covariances of the parameters by Monte Carlo techniques as suggested for instance by Singer and Cohen (1980). This involves generating say 1000 stochastic matrices from binomial samples of sizes $n_{1\cdot}$ and $n_{2\cdot}$, with parameters $\Pi_{ii} = n_{ii}/n_{i\cdot}$ $i = 1,2$, respectively. The binomial variates so produced are then used to compute the α's and β's using equation (34) by Newton Raphson or some other method. The Monte Carlo estimates (subscripted mc) are then obtained from the equations

$$\text{(i)} \quad E_{mc}(\hat{\alpha}/3) = 0.001 \sum_{j=1}^{1000} (\frac{\hat{\alpha}}{3})_j$$

$$\text{(ii)} \quad E_{mc}(\hat{\beta}) = 0.001 \sum_{j=1}^{1000} (\hat{\beta})_j$$

$$\text{(iii)} \quad var_{mc}(\hat{\alpha}/3) = 0.001 \sum_{j=1}^{1000} (\hat{\alpha}/3)_j^2 - (E_{mc}(\hat{\alpha}/3))^2 \qquad (36)$$

$$\text{(iv)} \quad var_{mc}(\hat{\beta}) = 0.001 \sum_{j=1}^{1000} (\hat{\beta})_j^2 - (E_{mc}(\hat{\beta}))^2$$

$$\text{(v)} \quad cov_{mc}(\hat{\alpha}/3, \hat{\beta}) = 0.001 \sum_{j=1}^{1000} (\hat{\beta})_j (\hat{\alpha}/3)_j - E_{mc}(\hat{\beta}) E_{mc}(\hat{\alpha}/3)$$

where the subscript j denotes a value of the respective parameter computed from the j-th matrix generated (cf. Singer et al. 1989). We used subroutine G05EDF from the NAG routine library (NAG, 1977) to generate these binomial variates for each of the 2 x 2 array entries of Table 3. The Monte Carlo estimates of the standard errors are the ones given in Table 6.

TABLE 6

Malaria Recovery and Incidence rates and their standard errors (s.e.) estimated by model C.

Survey pairs	Age groups (years)	Recovery rate (10^{-4}day^{-1}) $\hat{\alpha}_3$	(s.e.)	Incidence rate (10^{-4}day^{-1}) $\hat{\beta}$	(s.e.)	% correl. coeff.
(3,4)	0<1	37	(8)	33	(7)	11
	1-4	17	(1)	65	(15)	100*
	5-8	25	(2)	149	(24)	100*
	9-18	46	(5)	74	(9)	31
	19-28	81	(8)	38	(4)	31
	29-43	121	(10)	40	(3)	52
	44+	123	(13)	36	(4)	53
(4,5)	0<1	53	(10)	200	(22)	-18
	1-4	16	(1)	195	(27)	100*
	5-8	18	(1)	202	(28)	100*
	9-18	28	(2)	110	(10)	44
	19-28	69	(7)	65	(5)	44
	29-43	94	(8)	68	(5)	63
	44+	110	(16)	77	(9)	79
(5,6)	0<1	24	(6)	199	(40)	43
	1-4	15	(1)	158	(30)	100*
	5-8	24	(2)	170	(27)	100*
	9-18	42	(4)	98	(11)	48
	19-28	78	(7)	65	(6)	52
	29-43	106	(8)	57	(5)	66
	44+	117	(12)	54	(7)	72
(6,7)	0<1	24	(4)	44	(13)	27
	1-4	15	(1)	136	(28)	100*
	5-8	27	(2)	156	(26)	88
	9-18	50	(5)	74	(10)	42
	19-28	100	(8)	33	(4)	45
	29-43	124	(8)	27	(2)	56
	44+	135	(12)	27	(3)	61
(7,8)	0 1	36	(7)	13	(3)	5
	1-4	10	(1)	106	(24)	100*
	5-8	20	(1)	161	(24)	100*
	9-18	44	(5)	56	(8)	27
	19-28	85	(9)	33	(4)	32
	29-43	116	(9)	32	(2)	51
	44+	135	(17)	34	(4)	66

* Note the high correlation coefficients between the rates for 5-18 age groups.

While we defer, until a later section, the comparison of the estimates of the recovery and infection rates obtained by the various models, we need to note at this juncture that from tables 4-6, corresponding estimates of these rates decrease as r increases; that is, in the order of the models A, B and C. As these rates are bounded below by 0, it would seem that as $r \to \infty$ their values tend to some finite limit. However, we have not been able to obtain explicit expressions for these rates in terms of r; consequently we are not able to get analytically expressions for the limit. Nonetheless, we can obtain tight bounds for the rates if we consider an example of a model in which recovery time is assumed to be constant for each age group and survey pair.

6.3.3 A model with constant recovery times

This model, herein referred to as model D, may be thought of as having gamma distributed recovery times in which we let $r \to \infty$ and $\alpha \to 0$ with $r/\alpha = k$ fixed, to obtain a concentration of probability at k (Cox and Miller, 1977, page 258). In addition, we shall set a constraint on the parameter space by assuming that times between successive surveys do not exceed k. This has the result of restricting the number of an individual's transitions between successive surveys to at most one. The model is not therefore correct in the strict sense, for we need also to consider the case $k < t_{\ell}$. However, this case introduces the problem of deciding just how many transitions to allow between consecutive surveys. Cases have been recorded of uncomplicated P.falciparum infections clearing within two months (Kitchen, 1949), and we have also already mentioned the fact that some may last for as long as 1 year.

To derive the conditional probabilities $\phi_{ij}(t)$, we proceed as follows -

Let T_1 be the random time spent in state 1 before moving to state 2, measured from the time of the first survey, and exponentially distributed with mean $\frac{1}{\beta}$, Y, the random time spent in state 2 before moving to state 1, measured also from the first of the two surveys but distributed uniformly in (o,k). Let T be the random time spent in state 1, and like T_1, distributed exponentially with mean $1/\beta$. The time between the surveys is t_ℓ, $\ell = 1,2,\ldots,5$.

Then,

(i) $$\phi_{12}(t_\ell) = \text{Prob}\{T_1 < t_\ell\}$$

$$\int_o^{t_\ell} \beta e^{-\beta u}\, du = 1 - e^{-\beta t_\ell}$$

(ii) $$\phi_{21}(t_\ell) = \text{Prob}\{Y < t_\ell,\ Y + T > t_\ell\} \qquad (37)$$

$$= \int_o^{t_\ell} \text{Prob}\{Y + T > t_\ell \mid Y = y\}\, \frac{1}{k}\, dy$$

$$= \int_o^{t_\ell} \text{Prob}\{T > t_\ell - y\}\, \frac{1}{k}\, dy$$

$$= \frac{1}{k} \int_o^{t_\ell} e^{-\beta(t_\ell - y)} dy = (1 - e^{-\beta t_\ell})/\beta k$$

$$t_\ell \leq k$$

$$\ell = 1,2,\ldots,5.$$

As in previous models, we assume binomial sampling of the transition counts, and use maximum likelihood estimators of ϕ_{12} and ϕ_{21} (given in subsection 6.3.1) to obtain estimates

for β and k, from equations (37)

$$\text{(i)} \quad \hat{\beta} = -\log_e(1 - \hat{\phi}_{12})/t_\ell$$

$$\text{(ii)} \quad \hat{k} = \max \{ -t_\ell \hat{\phi}_{12} / (\hat{\phi}_{21} \log_e(1 - \hat{\phi}_{12})), t_\ell \} \tag{38}$$

provided $\hat{\phi}_{12}, \hat{\phi}_{21} \neq$ o or 1. These limitations are analogous to the embeddability criterion of subsection 2.2 which would restrict model D only to data with no zero transition counts such as those in Table 3.

With the conditions, $o < \hat{\phi}_{12}, \hat{\phi}_{21} < 1$, satisfied the asymptotic covariance matrix, $C(\hat{\beta},\hat{k})$, may again be obtained by the inverse of the Fisher information matrix, computed at $\beta = \hat{\beta}$ and $k = \hat{k}$.

$$C(\hat{\beta},\hat{k}) = - \begin{bmatrix} \frac{\partial^2 L}{\partial \beta^2} & \frac{\partial^2 L}{\partial \beta \partial k} \\ \frac{\partial^2 L}{\partial \beta \partial k} & \frac{\partial^2 L}{\partial k^2} \end{bmatrix}^{-1}_{\beta = \hat{\beta},\ k = \hat{k}} \tag{39}$$

where, as before, L is the logarithm of the likelihood function. The elements of the information matrix are given by

$$\text{(i)} \quad -\frac{\partial^2 L}{\partial \beta^2} = V^2 \{ S_1 U^2 \hat{\phi}_{12}^2 + S_2 \hat{\phi}_{21} (U + \hat{\phi}_{12})^2 \} / (U^2 \hat{\phi}_{12})^4$$

$$\text{(ii)} \quad -\frac{\partial^2 L}{\partial k^2} = S_2 U^2 \phi_{21}^4 / V^2 \tag{40}$$

$$\text{(iii)} \quad -\frac{\partial^2 L}{\partial \beta \partial k} = S_2 (U + \hat{\phi}_{12}) \hat{\phi}_{21}^3 / \hat{\phi}_{12}^2$$

$$k \geq t_\ell ,$$

where now we have

$$\text{(i)} \qquad U = (1 - \hat{\phi}_{12})\log_e(1 - \hat{\phi}_{12})$$

$$\text{(ii)} \qquad V = t_\ell \hat{\phi}_{12}(1 - \hat{\phi}_{12})$$

$$\text{(iii)} \qquad S_i = n_{i\cdot}/\{\hat{\phi}_{ij}(1-\hat{\phi}_{ij})\} \qquad i \neq j = 1,2.$$

From (40), the elements of matrix (39) become

$$\text{(i)} \quad \text{var}(\hat{\beta}) = \hat{\phi}_{12}/\{S_1 V^2\}$$

$$\text{(ii)} \quad \text{var}(\hat{k}) = V^2\{S_1 U^2 \hat{\phi}_{12}^2 + S_2 \hat{\phi}_{21}^2 (U+\hat{\phi}_{12})^2\} / \{S_1 S_2 \hat{\phi}_{12}^2 U^4 \hat{\phi}_{21}^4\} \qquad (41)$$

$$\text{(iii)} \quad \text{cov}(\hat{\beta},\hat{k}) = -(U+\hat{\phi}_{12})/\{S_1 \hat{\phi}_{21} U^2\}$$

$$k \geq t_\ell$$

The results of applying model D to each 2 x 2 array entry of Table 3 are listed in Table 7. Inspections of tables 4,5,6 and 7 show a consistent and gradual fall in the estimated rates in the order of the models A, B, C and D. That estimates of the rates from model A are higher than those from model D, can be shown to hold analytically.

Let $\hat{\alpha}_D = 1/\hat{k}$, where the subscript indicates the model used. From equations (19) and (38), suppose $\hat{k} > t_\ell$, then clearly

$$\frac{\hat{\alpha}_A}{\hat{\alpha}_D} = \frac{\hat{\beta}_A}{\hat{\beta}_D} = \frac{\dfrac{\log_e(1-\hat{\phi}_{12}-\hat{\phi}_{21})}{(\hat{\phi}_{12}+\hat{\phi}_{21})}}{\dfrac{\log_e(1-\hat{\phi}_{12})}{\hat{\phi}_{12}}}$$

TABLE 7

Malaria mean Recovery times and Incidence rates and their standard errors (s.e.), estimated by model D, plus comparisons with the results of model A

Survey pairs	Age groups (years)	Recovery time (days) $\hat{k}$	(s.e.)	Incidence rates (10^{-4} day^{-1}) $\hat{\beta}$	(s.e.)	% Correl. coeff.	Comparison with model A $\hat{\alpha}_{A-1}/\hat{\alpha}_D$	$\hat{\beta}_{A-1}/\hat{\beta}_D$
(3,4)	0<1	275	(70)	32	(8)	-11	0.15	0.19
	1-4	591	(87)	65	(14)	-30	0.06	0.08
	5-8	400	(56)	148	(23)	-46	0.16	0.16
	9-18	221	(21)	72	(9)	-28	0.26	0.25
	19-28	135	(11)	35	(3)	-14	0.45	0.43
	29-43	101	(5)	33	(2)	-15	0.81	0.79
	44+	100	(6)	29	(3)	-15	0.82	0.86
(4,5)	0<1	208	(78)	182	(24)	-20	0.77	0.77
	1-4	635	(120)	194	(25)	-39	0.14	0.13
	5-8	559	(92)	201	(26)	-45	0.17	0.16
	9-18	357	(42)	108	(10)	-27	0.18	0.19
	19-28	159	(13)	59	(4)	-19	0.51	0.51
	29-43	128	(7)	57	(3)	-19	0.80	0.79
	44+	117	(8)	60	(4)	-20	1.08	1.08
(5,6)	0<1	420	(170)	196	(39)	-29	0.26	0.25
	1-4	672	(128)	157	(28)	-47	0.14	0.11
	5-8	423	(59)	168	(25)	-57	0.23	0.21
	9-18	250	(22)	94	(10)	-41	0.32	0.31
	19-28	147	(9)	57	(5)	-22	0.65	0.65
	29-43	119	(5)	45	(3)	-27	0.98	0.98
	44+	113	(5)	41	(3)	-28	1.14	1.15
(6,7)	0<1	426	(115)	44	(13)	-18	0.11	0.11
	1-4	672	(127)	136	(26)	-44	0.08	0.10
	5-8	373	(57)	154	(23)	-53	0.19	0.21
	9-18	211	(19)	71	(9)	-36	0.33	0.31
	19-28	118	(6)	27	(3)	-25	0.71	0.74
	29-43	103	(3)	21	(2)	-22	0.97	1.00
	44+	99	(4)	20	(2)	-21	1.09	1.10
(7,8)	0<1	279	(59)	13	(5)	- 9	0.14	0.15
	1-4	959	(214)	106	(23)	-31	0.05	0.05
	5-8	488	(76)	161	(23)	-42	0.12	0.12
	9-18	231	(23)	55	(7)	-24	0.25	0.29
	18-28	130	(10)	29	(3)	-14	0.50	0.52
	29-43	104	(5)	26	(2)	-14	0.78	0.77
	44+	95	(6)	27	(3)	-14	0.99	0.96

$$= \frac{\hat{\phi}_{12}\log_e(1-\hat{\phi}_{12}-\hat{\phi}_{21})}{(\hat{\phi}_{12}+\hat{\phi}_{21})\log_e(1-\hat{\phi}_{12})} \qquad (42)$$

Since

$$-\frac{\log_e(1-\hat{\phi}_{12}-\hat{\phi}_{21})}{\hat{\phi}_{12}+\hat{\phi}_{21}} > -\frac{\log_e(1-\hat{\phi}_{12})}{\hat{\phi}_{12}}$$

then

$$\frac{-\dfrac{\log_e(1-\hat{\phi}_{12}-\hat{\phi}_{21})}{\hat{\phi}_{12}+\hat{\phi}_{21}}}{-\dfrac{\log_e(1-\hat{\phi}_{12})}{\hat{\phi}_{12}}} > 1 \qquad (43)$$

which implies $\dfrac{\hat{\alpha}_A}{\hat{\alpha}_D} = \dfrac{\hat{\beta}_A}{\hat{\beta}_D} > 1.$

Similar results are obtained when $\hat{\alpha}_D = \frac{1}{t_\ell}$.

While (43) still holds for $\dfrac{\hat{\beta}_A}{\hat{\beta}_D}$, we now have

$$\frac{\hat{\alpha}_A}{\hat{\alpha}_D} = -\frac{\hat{\phi}_{21}\log_e(1-\hat{\phi}_{12}-\hat{\phi}_{21})}{\hat{\phi}_{12}+\hat{\phi}_{21}}$$

$$> -\frac{\hat{\phi}_{21}\log_e(1-\hat{\phi}_{12})}{\hat{\phi}_{12}} \qquad (44)$$

But since from (38 (ii)) we have (because $\hat{k} = t_\ell$)

$$t_\ell \geq -t_\ell\hat{\phi}_{12} / \{\hat{\phi}_{21}\log_e(1-\hat{\phi}_{12})\},$$

then,

$$-\frac{\hat{\phi}_{21}}{\hat{\phi}_{12}} \leq \frac{1}{\log_e(1-\hat{\phi}_{12})},$$

from which we obtain the inequality

$$-\frac{\hat{\phi}_{21}\log_e(1-\hat{\phi}_{12})}{\hat{\phi}_{12}} \geq 1 \qquad (45)$$

A possible explanation for the low estimates of the rates from model D compared with those from A is that, model A allows for several transitions between consecutive surveys while model D is restricted to at most one transition only. It could be said that model D under-estimates the rates.

From tables 4 and 7, we observe that the discrepancies between the estimates of the rates by models A and D are greatest among the older age groups. These are the groups with generally high recovery rates most of which have their reciprocals (expected recovery times as by model A) less than the times between surveys. The quantities $\frac{\hat{\alpha}_A}{\hat{\alpha}_D} - 1$ or $\frac{\hat{\beta}_A}{\hat{\beta}_D} - 1$ may be used to assess the levels of discrepancies between corresponding estimates. These are given in the last two columns of Table 7. Among the younger age groups (0 - 8), these values are largest during the wet season (survey pairs (4,5) and (5,6)), the period of high infection rates and low recovery rates. Overall, these quantities are large for high rates and small for low rates. Despite these discrepancies, all the four models give qualititively similar patterns of results with respect to age and season.

We note in passing that model D may be used to describe data that are inconsistent with class of continuous-time Markov models. Some of the transition counts in the intervention and post-intervention phase of the Garki project are

of this type. For example, transition counts of infants in certain villages in Garki, at surveys 21 and 22 are given by Singer and Cohen as the matrix,

$$(n_{ij}) = \begin{bmatrix} 68 & 28 \\ 17 & 4 \end{bmatrix}$$

Since det $\underline{\Phi}$ has the same sign as det (n_{ij}) and det (n_{ij}) = 68 x 4 - 28 x 17 < o, then the matrix does not satisfy the embeddability condition. Now if we use equations (38) and (41), with t_{22} = 70 days, we obtain $\hat{\beta}$ = 0.0049 (s.e.= 0.0009) and $\hat{k}$ = 73.14 days (s.e = 0.56).

However, the restriction that the intervals between successive surveys should not exceed k may not always be satisfied. For example, if intervention includes mass drug administration of the de facto population, then it is possible that k may lie outside the restricted parameter space. This is because mass drug administration would bring about quick clearance of parasitamia and that means short recovery periods.

It is not possible to discriminate among these four models (A,B,C and D). Essentially, the first three may be considered as versions of the same model, with r = 1,2 and 3, respectively. Ideally, we would like to estimate r simultaneously with α and β. However, this would introduce the problem of non-identifiability of r and α(Silvey, 1975, page 50). Moreover, we cannot use the usual goodness of fit test since all the models fit the data equally; the log likelihood value for any 2 x 2 array is the same for any r. Alternatively, for each age group, we could simultaneously

estimate r, α and β across several 2 x 2 arrays. However, because of expected heterogeneity among these arrays with respect to season, it may not be appropriate to do so.

In view of the above difficulty, we may only speculate that the intermediate models (B and C) would provide reasonably good descriptions of the process. Model D may be useful to give simple and rough descriptions of non-embeddable transition counts.

6.3.4 The Results of Model Applications

As already stated, the corresponding estimates of both the (apparent) incidence and (apparent) recovery rates obtained from models A, B, C and D, exhibit similar qualitative patterns. These patterns are associated both with season and age. Figures 12 and 13 give visual summaries of the rates as computed by the four models. We briefly describe below the variations in these rates.

(a) Incidence Rates

From Figure 12 we see that incidence rates are generally low during the dry season ((i), (iv), (v)) and high during the wet season ((ii), (iii)). With regard to age, during the dry season the rates rise steadily to their highest level at the 5-8 years age group. They then decline, steeply at first, then gradually, to level off from age 2$. During the wet season, infants show the highest incidence rates which decline gradually at first, and then more rapidly from 5-8 years age group to 9-18 years age group. After that they level off, staying above the dry season level all the time. There is also a slight dip for the 1-4 years age group.

Fig 12 DAILY INCIDENCE RATES OF P.FALCIPARUM MALARIA BY AGE AND BY SEASON, ESTIMATED BY MODELS A, B, C AND D.

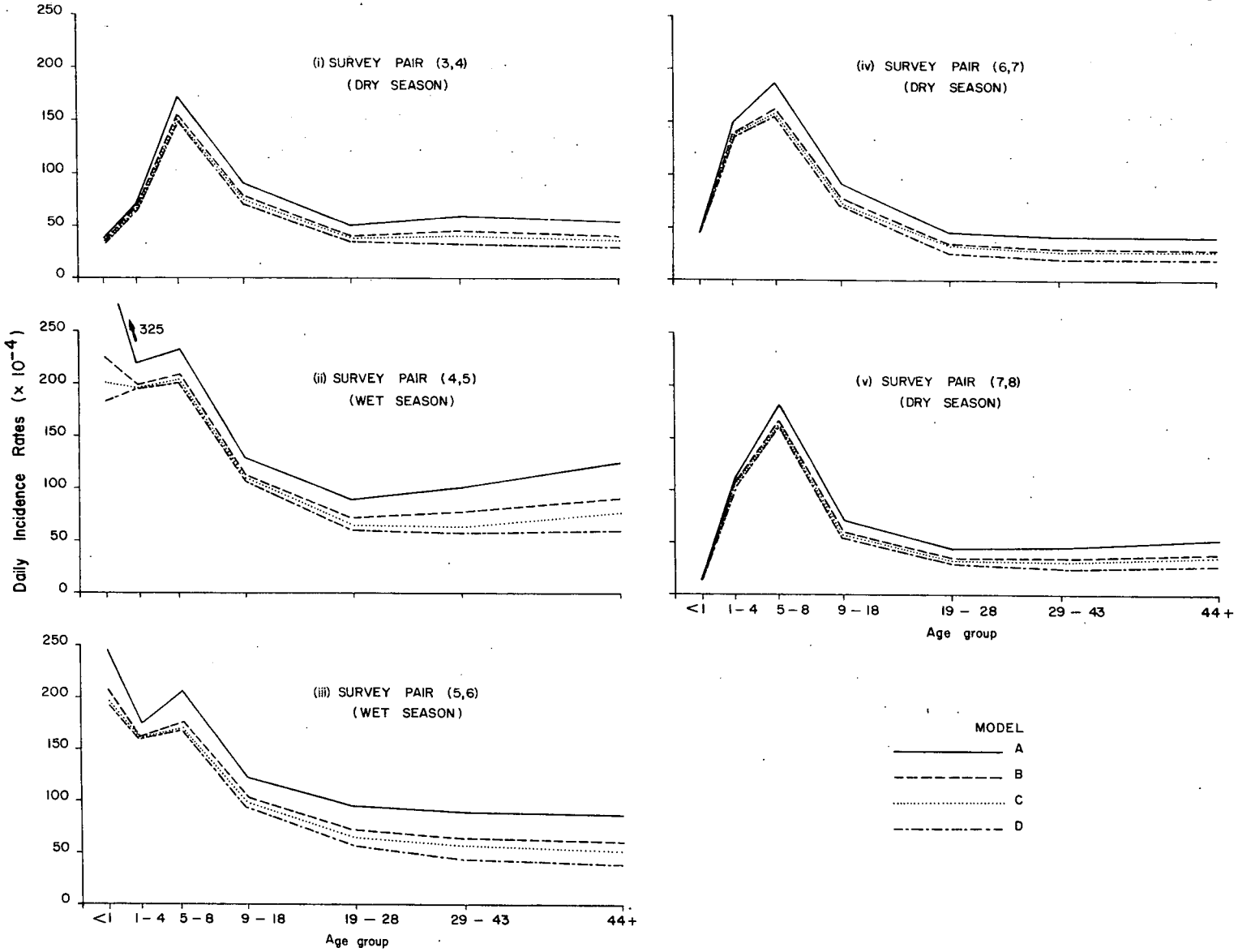
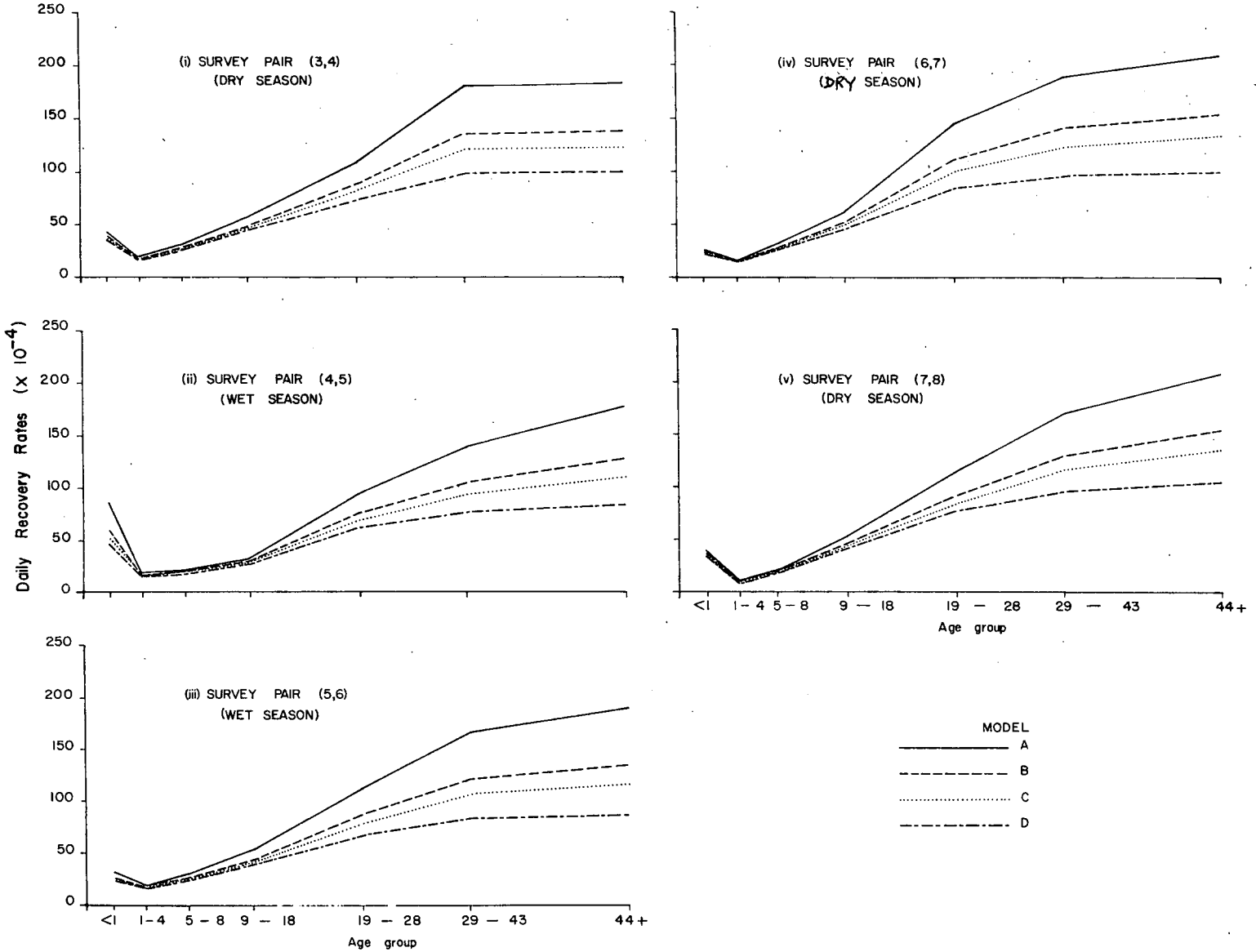


Fig 13 DAILY RECOVERY RATES FOR P.FALCIPARUM MALARIA BY AGE AND SEASON, ESTIMATED BY MODELS A, B, C AND D

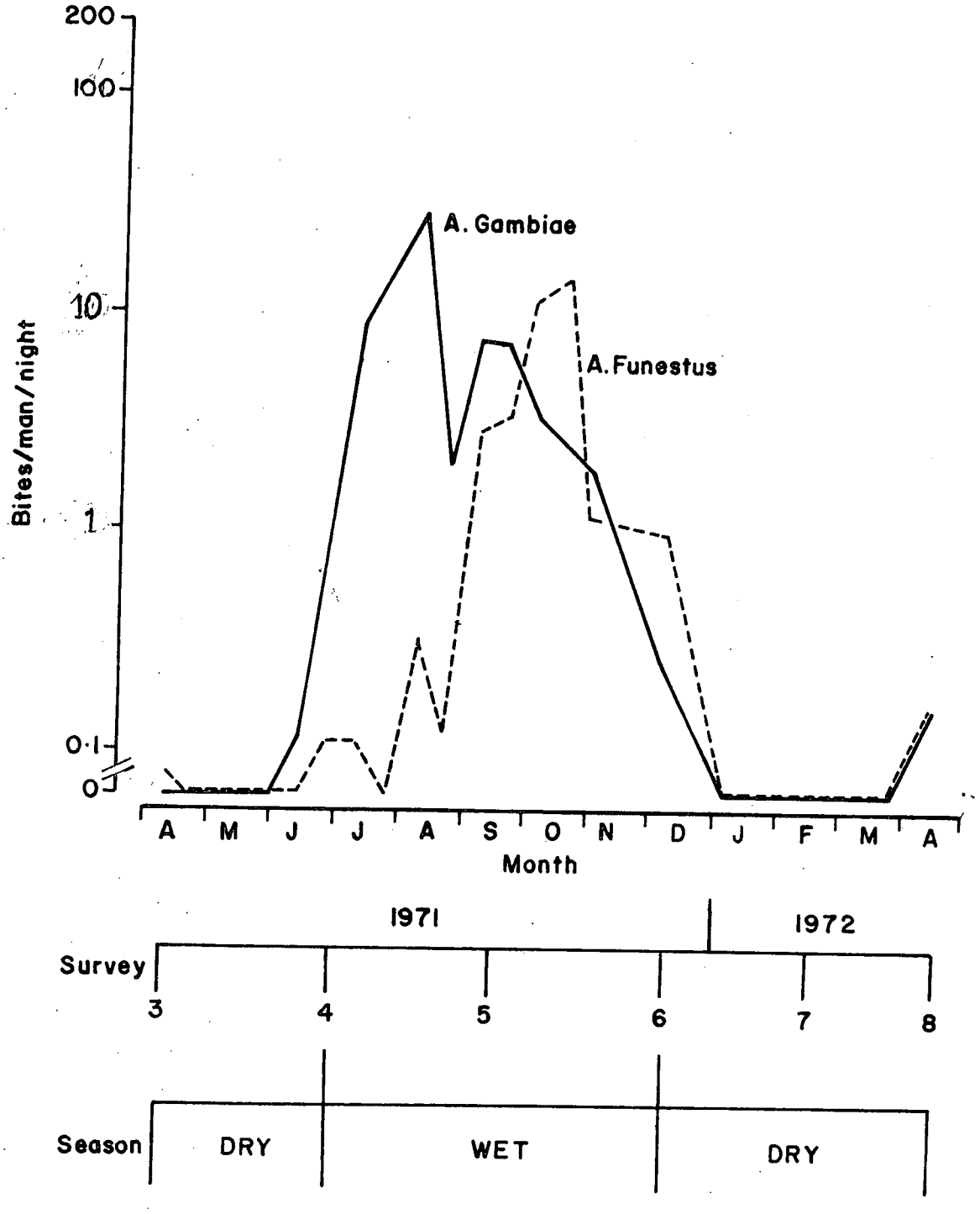


These age variations are usually interpreted in terms of immunity, and those by season are attributed to the variation of vector man-biting rates. Figure 14 reproduced from Molineaux et al. (1980, page 57), is a graph of man-biting rates of the ano pheline (A.) species A.gambiaes s.e. and A. funestus, from one of the villages included in the baseline data. A cursory comparison of this figure with Figure 12 indicates high incidence rates correspond to high vector man-biting rates. During the dry season when the man-biting rates are low, incidence rates should therefore result mainly from relapses.

Generally, individuals aged 1 - 18 years have high incidence rates and low recovery rates. This is because of the erosion by time of natural immunity and the fact that acquired immunity is not yet well established. This point is well illustrated in Figure 15. Here, immunoglobulin G (IgG) and immunoglobulin M (IgM) titres of individuals, expressed as percentages of maximum concentrations (for IgG, 100% = 520 IU/ML, and for IgM, 100% = 915 IU/ML (Molineaux et al., 1980)), are plotted against age. As mentioned earlier (Section 1.5 and Section 4.2), these antigens are long established as means to gauge immunity responses in individuals.

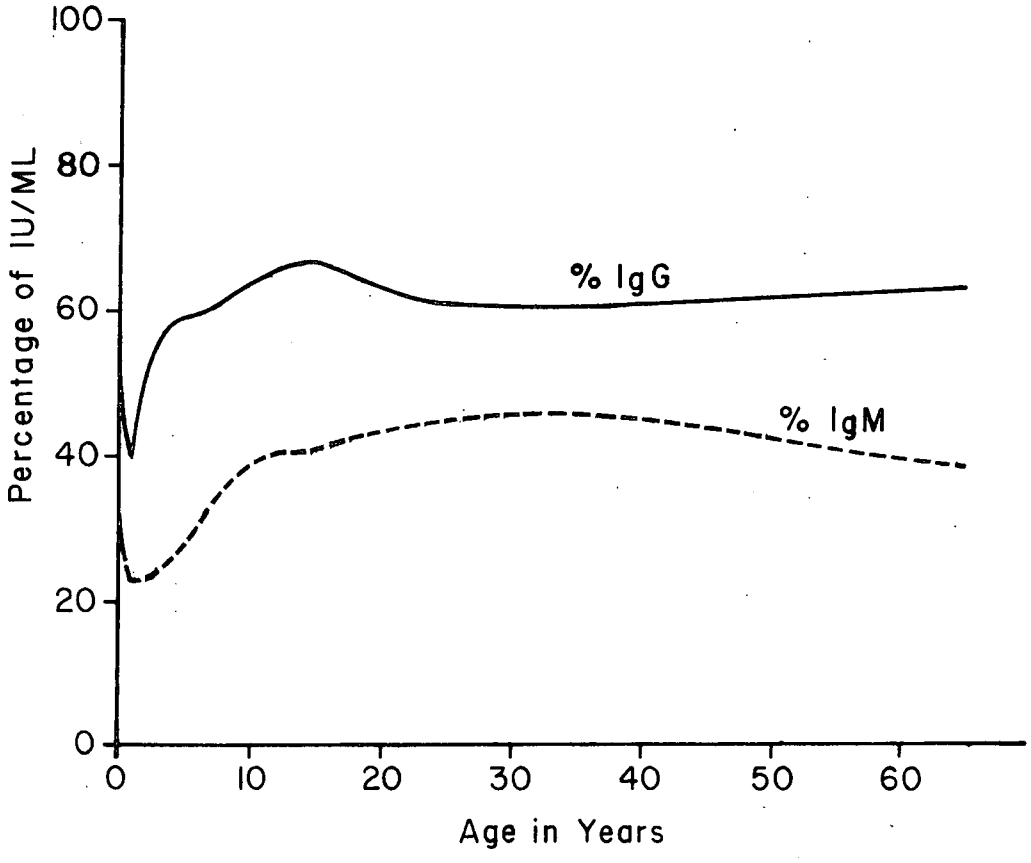
The serological data used to produce the graph cover only some of the people surveyed in survey 5 (wet season), and results are meant to be exploratory. The data have large variation, and to smooth them we used a robust locally weighted regression procedure (Cleveland, 1979). The procedure, which we outline in the Appendix, consists of local iterative fitting of polynomials to data using weighted least squares. The aim is to guard against deviant points

Fig 14. MAN-BITING RATES OF A.GAMBIAE s.l and A.FUNESTUS, ESTIMATED FROM NIGHT-BITE COLLECTIONS IN A TYPICAL UNTREATED VILLAGE IN GARKI



Source: Molineaux and Gramiccia, 1980 page 57

Fig. 15 PERCENTAGE LEVELS OF IgG AND IgM TITRE, BY AGE FOR ALL INDIVIDUALS IN SURVEY 5 (WET SEASON)



Note : For IgG 100% = 520 IU/ML
For IgM 100% = 915 IU/ML

that might distort the smoothing.

The rapid loss of natural immunity is clearly discernible from the figure (15). Both levels of IgG and IgM drop steeply to age one before they rise steadily to level off around the age of 15 years for IgM and, for IgG, three years later (18 years point).

The very high incidence rates in infants, during the wet season could imply, among other things that while natural immunity may enhance recovery (see next paragraph), it offers little or no protection against infection. The low incidence rates in the older age groups may be indicative of acquired immunity faring better in this respect.

We may note in passing that the incidence rates for all age groups are consistently higher earlier in the wet season than later. This is probably because the disease usually takes its toll at the beginning of the high transmission season, resulting in a drop in the pool of susceptibles for the rest of the wet season.

(b) Recovery Rates

Recovery rates exhibit patterns of variation opposite to those of incidence rates, both by season and age. Broadly, these rates are low (as opposed to high incidence rates) during the wet season (Figure 13, (ii) and (iii)), and high during the dry season (Figure 13, (i), (iv) and (v)).

In infants, recovery rate is lowest at the onset of the dry season (survey pair (6,7)) while children aged 1-4 have the lowest rate about the middle of the dry season

(survey pair (7,8)). Maximum rates are attained at the beginning of the wet season (survey pairs (4,5)). Individuals over the age of 8 years have the lowest recovery rate about the same time, and the maximum rate, early in the dry season (survey pair (6,7)). These variations again reflect the intensity of vector man-biting rates and relapses, especially among the older age groups.

We now briefly consider variations with age. Except for infants, recovery rates rise rapidly from 1-4 age group to 19-28 age group, and then stabilizes thereafter. Figures 13 and 15, taken together, illustrate that increase in recovery rates correspond with the rise of acquired immunity. The decline observed in the recovery rates from infants to 1-4 age group is indicative of the loss of natural immunity. It could also be partly due to the effect of superinfection, which may have had sufficient time to set in by then. However, during the dry season this factor is unlikely to be of significance.

6.3.5 Pooling rates by season

In view of the very low vector man-biting rates during the dry season (Figure 14), it may be reasonable to expect that the incidence rates estimated for that period of the year are largely due to relapse of previous infections. We also note that, except for children aged under 5 years, incidence (and also recovery) rates, when grouped by season (dry and wet), are of similar magnitude for the different age groups. This suggests we may use simple exploratory techniques to obtain, for each age group, the mean rates by season.

Columns 1 and 2 of Table 8 give the weighted means of the incidence rates for the dry and wet seasons respectively, (model A). Similarly, columns 4 and 5 are the corresponding weighted means of the recovery rates. The ratios of the rates appear in columns 3 (incidence rates) and 6 (recovery rates). The weights used here are the inverses of the corresponding asymptotic variances of the estimated rates. For instance, columns 1 and 4 are the results of the rates obtained from the survey pairs (3,4), (6,7) and (7,8) (dry season).

A visual display of Table 8 is provided by Figure 16. Again we note the high incidence rates in the wet season in infants, which drop steadily to 19-28 age group to stabilize thereafter. The very low incidence rate for infants in the dry season re-inforces the argument that the dry season infections are largely relapses of old infections. The incidence rate for infants are mainly due to new infections. The weighted recovery rates exhibit roughly the same patterns of variation as before (cf. Figure 13).

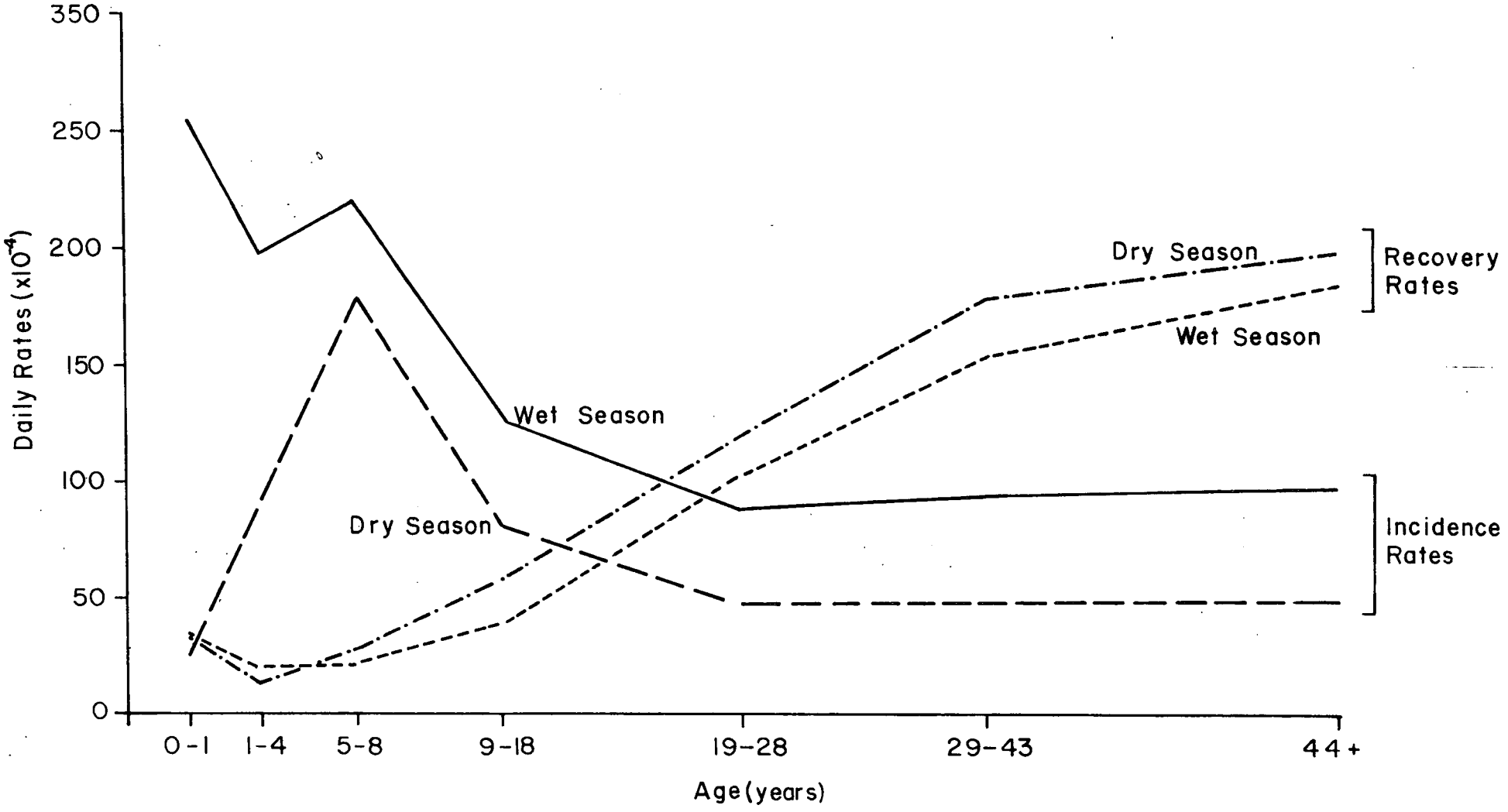
Columns 3 and 6 of Table 8 show a wide variation in incidence rates among younger age groups (0-4) and only small variations in recovery rates for all age groups. Except for 1-4 age group, recovery rates in the dry season are consistently higher than those in the wet season. The differences are more pronounced in the older age groups than in the younger ones. This could be interpreted to mean that either superinfection occurs more among adults than the young, or that relapses are cleared more easily than fresh infections. We consider the latter the more likely explanation.

TABLE 8

Weighted daily incidence and recovery rates (model A) for the dry and wet seasons

Rate	incidence rate, $\hat{\beta}$ (day^{-1})			recovery rate, $\hat{\alpha}$ (day^{-1})		
Season	dry	wet		dry	wet	
Age Group	(1)	(2)	(2)/(1)	(4)	(5)	(5)/(4)
0 < 1	0.0024	0.0255	10.63	0.0034	0.0032	0.94
1 - 4	0.0092	0.0198	2.15	0.0014	0.0017	1.21
5 - 8	0.0179	0.0219	1.22	0.0027	0.0024	0.89
9 - 18	0.0082	0.0127	1.55	0.0058	0.0040	0.69
19 - 28	0.0047	0.0091	1.94	0.0122	0.0103	0.84
29 - 43	0.0048	0.0095	1.98	0.0181	0.0154	0.85
44+	0.0049	0.0099	2.02	0.0200	0.0186	0.93

Fig. 16 WEIGHTED DAILY INCIDENCE AND RECOVERY RATES FOR WET AND DRY SEASONS (MODEL A)



It should be emphasized that these weighted estimates of the rates are at best exploratory. Nonetheless, as noted earlier, they do illustrate much more clearly, the difference between the effects of natural and acquired immunity on incidence and recovery rates. As indicated by the very high incidence rate of infants in the wet season, natural immunity seems to have virtually no effect on preventing infection occurring. Acquired immunity, on the other hand, appears first to enhance recovery then to reduce the occurrence of infection. The upward trend of recovery rate commences early, from 1-4 years age group, the downward trend of incidence rate, one age group later (5-8).

6.4 Inhomogeneous Markov Chain models

A natural extension to model A, in the last section, is to suppose that the parameters α and β are not constant between successive surveys. So let us assume in this section that the data may be described by a class of inhomogeneous continuous time Markov models. As mentioned earlier (subsection 2.2), the idea is to use suitably chosen forms for the elements of the intensity matrix $A.(\cdot)$ or hazard functions $\alpha(\cdot)$ and $\beta(\cdot)$, based on some specific features of the disease, provided these functions satisfy, say, the boundedness conditions given by (5).

6.4.1 Seasonally Dependent Infection and Recovery Rates

One of the main factors influencing the prevalence of the disease, which we can try to incorporate is seasonality, More specifically, if we assume homogeneity within each age

group (as in the previous models), seasonality may be explicitly included by supposing, for example, that the hazard functions vary sinusoidally with the time of the year as was suggested in Chapter 4 (for infection rate only). Thus suppose, for each age group,

$$\text{(i)} \quad \alpha(t) = a_1(1+a_2 \sin\omega t + a_3 \cos\omega t)$$

$$\text{(ii)} \quad \beta(t) = b_1(1+b_2 \sin \omega t + b_3 \cos\omega t) \tag{46}$$

where $\omega = 2\Pi/y$, a_i, b_i, $i = 1,2,3$, are constants with the constraints

$$0 < a_1, b_1 < 1$$

and

$$-1 < a_i, b_i < 1 \qquad i = 2,3$$

As before, (subsection 6.2.2) y is the period of the intensity matrix of the length of an epidemiological year, defined as the period between two successive minima of disease prevalence. This period could also be made one of the parameters to be estimated, but here, it is assumed to equal one calendar year. The parameters a_1 and b_1 are, respectively, the average rates of recovery and infection, while the a_i, b_i, $i = 2,3$ determine both the phases and relative amplitudes of the sinusoidal modulations.

By substituting (46) into (7) we obtain expressions for the transition probabilities $p_{ij}(s,t)$, $i,j = 1,2$, in terms of a_m, b_m, $m = 1,2,3$.

First, we note that

$$\int_s^t (\alpha(u)+\beta(u))\,du = \int_s^t \{ (a_1+b_1) + (a_1a_2+b_1b_2)\sin\omega u + (a_1a_3+b_1b_3)\cos\omega u \}\,du$$

$$= (a_1+b_1)(t-s) - \left(\frac{a_1a_2+b_1b_2}{\omega}\right)(\cos\omega t-\cos\omega s) + \left(\frac{a_1a_3+b_1b_3}{\omega}\right)(\sin\omega t-\sin\omega s) \qquad (47)$$

After appropriate substitution of (46) and (47) into equation (7 (i)), we have

$$p_{12}(s,t) = b_1 \int_s^t (1+b_2\sin\omega v+b_3\cos\omega v) \exp\{ -(a_1+b_1)(t-v) + \left(\frac{a_1a_2+b_1b_2}{\omega}\right)(\cos\omega t-\cos\omega v) - \left(\frac{a_1a_3+b_1b_3}{\omega}\right)(\sin\omega t-\sin\omega v) \}\,dv$$

$$= b_1 h(t) \int_s^t (1+b_2\sin\omega v+b_3\cos\omega v)\, g(v)\,dv, \qquad (48)$$

where

(i) $h(t) = \exp\{ -(a_1+b_1)t + \left(\frac{a_1a_2+b_1b_2}{\omega}\right)\cos\omega t - \left(\frac{a_1a_3+b_1b_3}{\omega}\right)\sin\omega t \}$;

(ii) $g(v) = \exp\{ (a_1+b_1)v - \left(\frac{a_1a_2+b_1b_2}{\omega}\right)\cos\omega v + \left(\frac{a_1a_3+b_1b_3}{\omega}\right)\sin\omega v \}$.

Now, since

$$\int_s^t g(v)dv = \frac{g(t)-g(s)}{a_1+b_1} + \frac{\omega}{a_1+b_1}\int_s^t (c_1 \sin\omega v - c_2 \cos\omega v)g(v)dv, \quad (49)$$

where

$$\text{(i)} \quad c_1 = \frac{a_1 a_2 + b_1 b_2}{\omega}$$

$$\text{(ii)} \quad c_2 = \frac{a_1 a_3 + b_1 b_3}{\omega},$$

equation (46) becomes

$$p_{12}(s,t) = \frac{b_1(1-h(t)g(s))}{a_1+b_1} + \frac{a_1 b_1 h(t)}{a_1+b_1}\{(b_2-a_2)\int_s^t \sin\omega v g(v)dv + (b_3-a_3)\int_s^t \cos\omega v g(v)dv\} \quad (50)$$

Similarly. we have

$$p_{21}(s,t) = \frac{a_1(1-h(t)g(s))}{a_1+b_1} + \frac{a_1 b_1 h(t)}{a_1+b_1}\{(a_2-b_2)\int_s^t \sin\omega v g(v)dv + (a_3-b_3)\int_s^t \cos\omega v g(v)dv\} \quad (51)$$

We note that the integrals in (50) and (51), which are essentially sine and cosine transforms of g(v), are not amenable to analytic integration. Some bounds can be obtained for the $p_{ij}(s,t)$, but these depend too much on the signs and relative magnitudes of the a_i and b_i, i = 1,2,3, to be of use.

6.4.2 Fitting Model E to data

We shall apply this model, to be called model E, to two separate sets of data. The first set is that in Table 3. The other set which appears in Table 9 is that of individuals who have been present in all six baseline surveys. The parameters are estimated for each data set by age group over the six surveys which stretch roughly over one epidemiological year. Again we use maximum likelihood estimation. Unlike for models A, B, C and D, the log likelihood functions are summed over all the six surveys.

Let $o = t_1 < t_2 < \cdots < t_6$ denote the points in time at which the surveys were conducted. The time of the first survey is taken as the initial time zero. Further, let $n_{ij\ell}$ $i,j = 1,2$, denote the transition counts of individuals in state j at survey ℓ given they were in state i at survey $\ell-1$, $\ell = 3,2\ldots,8$. The log likelihood function

$$L = \sum_{\ell=3}^{8} \sum_{i,j=1}^{2} n_{ij\ell} \log_e p_{ij}(t_{\ell-i}, t_\ell) + \text{a constant term} \qquad (52)$$

was maximized for each age group, using subroutine E04JBF from the NAG routine library (NAG,1977). Another subroutine, D01ANF, also from the NAG library, was used to compute approximations for the integral expressions in (50) and (51). The asymptotic covariance matrix was obtained again by inverting the Fisher information matrix of (52). The results of the computation for both data sets are given in Table (10). Because some of the 2 x 2 arrays for infants, in the second data set, had zero elements, we have combined the o<1 age group data with those of 1-4 year age group.

TABLE 9

Survey Pairs	AGE GROUP											
	0<1		1-4		5-8		9-18		19-28		29-43		44+	
(3,4)	9	3	49	18	22	38	75	39	241	67	690	152	422	92
Dry season	1	0	39	279	40	385	57	171	65	62	122	72	88	42
(4,5)	1	9	23	65	13	49	63	69	208	98	581	231	340	170
Wet season	1	2	24	273	40	383	32	178	58	71	117	107	75	59
(5,6)	1	1	9	38	17	36	47	48	179	87	504	194	307	108
Wet season	1	10	29	309	59	373	59	188	74	95	209	129	149	80
(6,7)	2	0	12	26	22	54	61	45	216	37	612	101	395	61
Dry season	0	11	32	315	49	360	75	161	112	70	226	97	138	50
(7,8)	2	0	20	24	21	50	97	39	267	61	698	140	446	87
Dry season	1	10	25	316	49	365	62	144	49	58	123	75	76	35

Malaria Parasitology Transition Data : Observed Numbers n_{ij} of persons in state i at one survey and state j at the next survey, for individuals present in all surveys.

TABLE 10

Age groups (years)	Data Set	Parameters							
		Recovery				Infection			
		a_1	a_2	a_3	a'	b_1	b_2	b_3	b'
0<1	I	0.0036	0.2018	0.3392	0.3947	0.0111	0.3695	-0.8812	0.9555
	II	-	-	-	-	-	-	-	-
1-4	I	0.0017	0.1608	-0.1645	0.2300	0.0154	0.0323	-0.5117	0.5127
	II*	0.0024	0.0367	-0.2789	0.2813	0.1740	-0.1756	-0.7115	0.7264
5-8	I	0.0027	-0.1393	0.0136	0.1400	0.0197	0.0785	-0.0988	0.1262
	II	0.0032	-0.1865	-0.0583	0.1954	0.0199	0.0085	-0.0196	0.0214
9-18	I	0.0051	-0.2368	0.1027	0.2581	0.0103	0.1595	-0.3000	0.3398
	II	0.0056	-0.3473	0.1349	0.3725	0.0096	0.0452	-0.3451	0.3480
19-28	I	0.0115	-0.2028	0.0720	0.2152	0.0067	0.1976	-0.3877	0.4352
	II	0.0122	-0.0948	0.1440	0.1724	0.0059	0.3181	-0.2244	0.3893
29-43	I	0.0169	-0.1404	0.1262	0.1888	0.0069	0.3113	-0.3146	0.4426
	II	0.0178	-0.1519	0.0821	0.1727	0.0061	0.1949	-0.2639	0.3280
44+	I	0.0193	-0.1787	0.1330	0.2228	0.0072	0.2931	-0.3111	0.4278
	II	0.0212	-0.1356	0.1451	0.1986	0.0069	0.2976	-0.2514	0.3896

Estimated parameters of time-dependent recovery and incidence rates, plus amplitudes (a', b'), from data sets of Table 3 (I) and Table 8 (II).

* The parameter values are for the combined age groups 0 - 4 years.

We have also included, in the table, the relative amplitudes of the rates about their respective means, given by $(a_2^2 + a_3^2)^{\frac{1}{2}}$ and $(b_2^2 + b_3^2)^{\frac{1}{2}}$ for the recovery and infection rates, respectively. But before we proceed to discuss the results in Table 10, we shall try to assess how well the model fits the data.

6.4.3 Goodness of fit of model E

We use the familiar Pearson's goodness of fit statistic, denoted by G^2, which in our case is

$$G^2 = \sum_{\ell=3}^{8} \sum_{i,j}^{2} (n_{ij\ell} - \hat{n}_{ij\ell})^2 / \hat{n}_{ij\ell} \quad ,$$

where, $\hat{n}_{ij\ell}$ are the expected values of $n_{ij\ell}$, estimated by model E. G^2 approximates a chi-squared distribution for large observations. Table 11 gives the G^2 values for each age group together with the separate contributions of the survey pairs in the group. There are 4 degrees of freedom.

The table (11) shows that, for the data from Table 3, the model fits rather poorly for three of the seven groups, namely 1-4 years and 29+ years age groups. The main discrepancies come from the end survey pairs (3,4) and (7,8). The intermediate age groups data fit quite well. With regard to data from Table 9, at 99% quantile point only one of the seven age groups, (29-43), has a poor fit. An inspection of the separate components of G^2 by survey pair again reveals the end survey pairs as the major source of the discrepancies, and the intermediate groups as fitting best. These rather large discrepancies from the end survey pairs are possibly

TABLE 11

Age groups (years)	Data Set	Survey pairs					
		(3,4)	(4,5)	(5,6)	(6,7)	(7,8)	G^2
0<1	I	2.795	3.921	0.535	3.339	0.595	11.186
	II	-	-	-	-	-	-
1-4	I	4.508	3.215	1.812	0.476	3.440	13.450†
	II	2.951	2.514	0.743	2.777	3.637	12.621*
5-8	I	3.441	1.549	0.352	0.143	2.248	7.733
	II	0.865	1.227	1.038	0.614	0.294	4.037
9-18	I	2.439	1.509	0.909	0.167	1.044	6.068
	II	0.513	0.398	0.480	0.249	0.341	1.981
19-28	I	1.263	0.123	1.016	2.858	3.341	8.600
	II	0.850	0.350	2.289	4.016	5.306	12.812
29-43	I	5.919	1.313	0.526	1.935	6.105	15.797†
	II	2.878	0.408	2.304	3.270	5.197	14.057†
44+	I	6.904	3.574	0.282	1.423	4.350	16.533†
	II	4.250	1.609	0.437	0.813	3.218	10.327

Goodness of fit of model E for transition frequencies of Table 3 (I) and Table 9 (II). d.f. = 4.

Rows with high values of G^2 are marked '†'

*The row is for the combined age groups 0 - 4 years

due to the symmetric forms of the rates. We have tried various non-symmetric forms for $\alpha(\cdot)$ and $\beta(\cdot)$ but none of them has improved on the results of Table 11.

6.5 Summary of results

Despite the poor fits in some of the age groups, the results of applying model E to both sets of data exhibit the effects of age, and hence immunity, in the estimated values of a_1 and b_1, the mean recovery and infection rates respectively.

The mean infection rate, $\hat{b}_1$, rose from 0.0111 (s.e. = 0.0014) among infants (o < 1 year) to a maximum of 0.0197 (s.e. = 0.0016) among children aged 5-8 years, and then dropped to stabilize around 0.0070 in individuals aged 19+ years. For the second data set, except for infants (now combined with 1-4 years age group), and 29-43 years age group, the estimates of b_1 are not significantly different from those of the first data set (Table 10).

The estimates of the mean recovery rate, a_1, from the first data set fell at first from 0.0036 (s.e. = 0.0006) in infants to 0.0017 (s.e. = 0.0002) in children aged 1-4 years, then rose steadily to 0.0193 (s.e. = 0.0012) in the 44+ age group. For the second data set, it rose from 0.0024 (s.e. = 0.0003), gradually at first, up to age group 9-18, then rapidly to 0.0212 (s.e. = 0.0016) in age group 44+. Again the corresponding estimates of a_1 in the two data sets are not significantly different.

The columns denoted as b' and a' in Table 10 give the estimates of the amplitudes of the seasonal variations of the

rates about their respective means. For both data sets, the variations for the infection rates were much higher than those for recovery rates, except for age groups 5-8 (first data set) and 9-18 (second data set). This suggests that infection rates variation by season are more pronounced than those of recovery rates. The younger age groups (0-4) exhibited the largest variation, and the middle age groups the least variation. Except for age groups 0-4 and 9-18, the seasonal variations of the recovery rates were generally low. These results conform with those discussed in subsection 3.5 (cf. Figure 16 and Table 8).

Overall, the results obtained here are comparable qualitatively with those obtained from models A-D. Perhaps one advantage model E has over models A-D, is that, for a given age group we could estimate infection and recovery rates between two consecutive surveys. The phases of these rates could also be estimated as the angles whose tangents are b_2/b_3 and a_2/a_3 respectively.

CHAPTER SEVEN

CONCLUSION

We started by giving a critical review of the basic model of Ross and its extensions, in which we showed that the various models proposed for superinfection are in fact special cases of a queueing process in equilibrium. We then proposed another model which allows for diminishing chance of re-infection as the number of broods increases in an infected person. This model lies somewhere between those of Dietz and Ross, and in this respect is similar to a class of models proposed by Aron et al.

In practice, the phases of the progress of malaria infection in an individual are subject to random effects. Motivated by this idea, we tried to represent the course of the disease by a hybrid semi-Markov model. We defined four mutually exclusive statuses to which any person may belong : susceptible, latent, infectious and non-infectious (but infected). Each status was further sub-divided into discrete immunity levels. This framework allows for the incorporation of the host-parasite relation as random while some aspects of the host-vector relationship are included as deterministic.

The model does not yield usable results by analytic methods. But when we use simulation it gives results that are generally in agreement with reality. For instance, the results show that while it is good public health practice to reduce infection rate, it may rob the population of its resistance to future attacks, explaining the causes of

resurgence of the disease in areas where it has been previously eradicated. The use of vaccine is shown to be the more effective control.

Because of lack of data to validate the model as described in Chapter 4, we reduced it to a 2-state stochastic model. Individuals were grouped into either the 'apparent' susceptible status (state 1) or 'apparent' infected status (state 2). For practical purposes state 1 was defined as that state in which the parasites of malaria could not be microscopically detected. The model was then used to obtain maximum likelihood estimates of the 'apparent' infection and recovery rates from malaria survey data from the WHO Garki project.

Several versions of the model,one homogeneous and Markovian (model A) and the others, non-Markovian (models B, C and D), were used. All the corresponding estimates of the rates from models A - D exhibited qualititively similar patterns, reflecting some of the more important host-vector and host-parasite relationships. We could not, however, discriminate among these models because they fitted the data equally well.

As illustrated by figures 12 and 13, the estimates show that infection rates are high in the wet season and low in the dry season, corresponding to the high and low vector densities during the wet and dry seasons, respectively. The effect of superinfection, a consequence of high vector density, is exhibited by the generally low 'recovery' rates during the wet season and high rates during the dry season.

As regards the host-parasite interaction, this is clearly illustrated by the difference in rates by age. The very high infection rates among infants during the wet season suggest that natural immunity offers little or no protection against infections. However, it appears to accelerate the clearance of parasitamia once infection has occurred, hence the high recovery rates relative to the 1 - 8 year age groups. Apart from the infants, both the decrease in infection rates and increase in recovery rates, with increasing age, are due to the cumulative effects of acquired immune responses from attacks by malaria over the years.

Results similar to those just described were obtained when we used a 2-state inhomogeneous Markov model with the transition intensities expressed as some periodic functions of time. The rate parameters were estimated by the maximum likelihood method, with the longitudinal panel data summed across the surveys. Two sets of data were used, that of Table 3, in which some individuals missed at least one survey, and that of Table 9, where only persons who appeared in all six surveys were included.

The results of applying the above model (Table 10) show much more clearly that seasonal fluctuations are in general more pronounced in infection rates than in recovery rates. Except for the 1 - 18 year age group, the amplitudes of the rates of infection are consistently higher than their corresponding values of the recovery rates. In addition, the differences between corresponding estimates of the parameters for the two data sets, are relatively minor, with the estimates of Table 8 showing the higher mean rates.

The model did not, however, fit the data as well as models A - D. The discrepancies, confined mainly to the end survey pairs, indicate that non-symmetric periodic functions for the rates may be more appropriate.

For a proper validation of the model described in Chapter 4, it is necessary to have accurate estimates of the parameters we have used. This requires a better quantitative understanding of the phases of the disease through experiments of specific aspects of the disease.

In estimating the 'apparent' infection and recovery rates we have not taken account of the effects of mixed infections, that is, simultaneous presence in the human body of two or more different species of the parasite. The analysis of mixed infections from the Garki project has been reported by Molineaux et al. (1980). We could use the data of mixed infections to distinguish among the models of superinfection discussed in Chapter 2. This may be done by assuming that mixed infections approximate cases of superinfection. The differences in the species should, however, be noted.

As stated by Singer et al., the theory of the analysis of longitudinal panel data, of the type from the Garki project, is not well developed. Data from the intervention and post-intervention periods of the project may require simple descriptions by non-Markovian models. Moreover, there is evidence to suggest that even the transition frequencies from the base-line data may not be Markovian. The use of model D is restricted by its constrained parameter space. There is clearly a need for better and preferably simple non-Markovian models to analyse such data.

The estimations of rates have been carried out by age groups. The data from Garki also contain classifications of individuals by titre of IgG, IgM and other measures of immune responses. It may be useful to estimate the infection and recovery rates according to the levels of these antigens and compare the results with those obtained here by age groups.

Finally, it has been argued that for models of biological phenomenon, such as infectious diseases, to be of value, they should be constructed in conjunction with some on-going field studies. We give a qualified support for this view. Because of the high expenses usually required, the long time spans involved and the efforts that often go into assembling inter-disciplinary teams, such field experiments are very rare. For instance, in the case of malaria, the WHO Garki project is probably the only comprehensive study of malaria ever carried out. On the other hand some progress may be made by using results from small experiments about specific facets of the disease. This is by no means ideal, but it may provide the only avenue to build an integrated malaria model in future.

APPENDIX

Smoothing Scatterplots

The method used to smooth the serological data in Chapter 6 is the so-called robust locally weighted regression method (Cleveland, 1979, and the references therein). It combines the usual weighted regression method with a robust fitting procedure which we now describe.

Suppose that (x_i, y_i), $i = 1,2,\ldots,n$, are the points of a scatterplot. Let $o \leq f \leq 1$ and $r = \lfloor fn \rfloor$, the truncated integral part of fn. For each x_i, let $W_k(x_i)$ define the weights for all x_k, $k = 1,2,\ldots,n$. The following is the sequence of operation.

1. For each i, compute the estimates $\hat{\beta}_j(x_i)$, $j = o, 1,\ldots,d$ of parameters in a local polynomial regression of degree d of y_k on x_k, fitted by weighted least squares with weights $W_k(x_i)$ for (x_k, y_k). The smoothed point at x_i is $(x_i, \hat{y}_i)$, where

$$\hat{y}_i = \sum_{j=o}^{d} \hat{\beta}_j(x_i) x_i^j \qquad (1)$$

2. Define B, a bisquare weight function by

$$B(x) = \begin{cases} (1-x^2)^2 , & |x| < 1 \\ o , & |x| \geq 1, \end{cases} \qquad (2)$$

and let $e_i = y_i - \hat{y}_i$, the residuals from the current fitted values. If we denote by s the median of the $|e_i|$, then robustness weights are defined by

$$\delta_k = B(e_k/cs) , \qquad (3)$$

where c is some positive constant integer (recommended value of c = 6).

3. Compute new $\hat{y}_i$ for each i by fitting a d-th degree polynomial as in step 1 but with new weights $\delta_k w_k(x_i)$ for (x_k, y_k), k = 1,2,...,n.

4. Repeat steps 2 and 3, for a pre-set number of times t, to obtain the final robust locally weighted regression fitted values, $\hat{y}_i$, i = 1,2,...,n.

The weight function W(x) for generating $w_k(x_i)$ is a 'tricube';

$$W(x) = \begin{cases} (1 - |x|^3)^3 & , \quad |x| < 1 \\ 0 & , \quad |x| \geq 1 \end{cases} \qquad (4)$$

in which $w_k(x_i)$ decreased as the distance of x_k from x_i increases. For each i, let h_i be the distance from x_i to the r-th nearest neighbour of x_i. Then, for k = 1,2,...,N

$$w_k(x_i) = W((x_k - x_i)/h_i) \qquad (5)$$

The choice of f determines the amount of smoothing required. Suggested values range from 0.2 ('rough' smoothed plot) to 0.8 ('smooth' smoothed plot) with 0.5 as a good compromise. The choices of values for d and t are largely of computational interest, and setting d = 1 and t = 2 are usually quite adequate for most situations. The tricube form for W(x) is recommended because it is supposed to provide chi-squared distribution approximation of error variance (ibid). The form for B(x) is said to perform well for robust estimation and regression (Gross, 1976, 1977). Suggested refinements include the handling of multiple y values of the x_i.

REFERENCES

Anderson, R.M., May, R.M. (1979) Population biology of infectious diseases : I. Nature, 280 : 361-367.

Arfken, G. (1970) Mathematical Methods for Physicists (2nd edn.) Academic Press, New York.

Armitage, P. (1953) A note on the epidemiology of malaria. Tropical Diseases Bulletin, 50 : 890-892.

Aron, J.L., May, R.M. (1980) The population dynamics of malaria (to appear).

Bailey, N.T.J. (1975) The Mathematical Theory of Infectious Diseases: (2nd edn.) Griffin, London.

Bartlett, M.S. (1964) The relevance of stochastic models for large-scale epidemiological phenomena. Applied Statistics, 13 : 2-8.

Bartlett, M.S. (1966) Some notes on epidemic theory. In : Research Papers in Statistics. Festschrift for J. Neyman (ed. F.N. David), John Wileys & Sons, Inc., New York, pp.445-459.

Bekessy, A., Molineaux, L., Storey, J. (1976) Estimation of incidence and recovery rates of Plasmodium falciparum parasitamia from longitudinal data. Bulletin of WHO, 54 : 685-691.

Belding, D.L. (1965) Textbook of Parasitology (3rd edn.). Appleton-Century-Crofts, New York.

Benenson, A.S. (1975) Control of Communicable Diseases in Man (12th edn.). American Public Health Association, New York.

Billingsley, P. (1979) Probability and Measure (1st edn.) John Wileys and Sons, Inc., New York.

Boyd, M.F. (1949) Epidemiology of malaria : factors related to the intermediate host. In : Malariology (ed. M.F.Boyd). W.B. Saunders, Philadelphia, pp.551-607.

Bruce-Chwatt, L.J. (1965) Malaria research for malaria eradication. Transactions of the Royal Society of Tropical Medicine and Hygiene, 59(1) : 105-137.

Bruce-Chwatt, L.J. (1969) Quantitative epidemiology of Tropical Diseases. Transactions of the Royal Society of Tropical Medicine and Hygiene, 63(1) : 131-143.

Bruce-Chwatt, L.J. (1974) The Resurgence of malaria and its control. Journal of Tropical Medicine and Hygiene, 77(4): 62-66.

Bruce-Chwatt, L.J. (1975) Mathematical models in Epidemiology and control of malaria. Tropical and Geographical Medicine, 28 : 1-8.

Bruce-Chwatt, L.J. (1979) Man against malaria : Conquest or defeat. Transactions of the Royal Society of Tropical Medicine and Hygiene, 73 : 605-617.

Bruce-Chwatt, L.J. (1980) Essential Malaria, Heinemann Medical Publications, London.

Bruce-Chwatt, L.J; Black, R.; Canfield, C.J.; Clyde, D.F.; Peters, W.; Wernsdorfer, W.H. (1981) Chemotherapy of Malaria. WHO, Geneva.

Burnside, W.S.; Panton, A.W. (1928) The Theory of Equations. Vol.I. Dublic University Press Series.

Chadwick, P.R.; Invest, J.F.; Bowron, J.B. (1977) An example of cross-resistance to pyrethroids in DDT-resistant Aedes aegypti. Pesticide Science, 8 : 618-624.

Christophers, S.R. (1949) Endemic and Epidemic prevalence. In: Malariology (Ed. M.F.Boyd) W.B. Saunders, Philadelphia, pp.698-721.

Chung, K.L. (1967) Markov Chains with Stationary Transition Probabilities (2nd edn.) Springer-Verlag, Berlin.

Cinlar, E. (1969) Markov renewal theory. Advances in Applied Probability, 1 ; 123-187.

Cinlar, E. (1975) Markov renewal theory - a survey.Management Science, 21 : 727-752.

Cleveland, W.S. (1979) Robust locally weighted regression and smoothing scatterplots. Journal of the American Statistical Association, 74(368) : 829-836.

Cohen, J.E. (1973) Heterologous immunity in human malaria. Quarterly Review in Biology, 48 : 467-489.

Cohen, S. (1979) Review lecture. Immunity to malaria. Proceedings of the Royal Society, London, B, 203 : 323-345.

Collins, W.E.; Jeffery, G.M.; Skinner, J.C. (1964) Fluorescent antibody studies in human malaria. II. Development and persistence of antibodies to P. falciparum. American Journal of Tropical Medicine and Hygiene, 13 : 256-260.

Comins, H.N. (1977) The development of insecticide resistance in the presence of migration. Journal of Theoretical Biology, 64 : 177-197.

Cornille-Brögger, R., Molineaux, L., Mathews, H.M., Storey,J. (1978) Longitudinal serological study of malaria in infants in the West African Savannah. Bulletin of WHO, 56 : 573-578.

Cox, D.R., Miller, H.D. (1977) The Theory of Stochastic Processes, Chapman and Hall, London.

Curtis, C.F., Cook, L.M., Wood, R.J. (1978) Selection for and against insecticide resistance and possible methods of inhibiting the evolution of resistance in mosquitoes. Ecological Entomology, 3 : 273-287.

Dame, D.A., Lofgren, C.S., Ford, H.R., Boston, M.D., Baldwin, K.V., Jeffrey, G.M. (1974) Release of chemosterilized males for the control of Anopheles albimanus in El Salvador. 2. Methods of rearing, sterilization and distribution. American Journal of Tropical Medicine and Hygiene, 23 : 282-287.

Dietz, K., Molineaux, L., Thomas, A. (1974) A malaria model tested in the African Savannah. Bulletin of WHO, 50 : 347-357.

Dietz, K. (1976) The incidence of infectious disease under the influence of seasonal fluctuations. In : Mathematical Models in Medicine. (ed. J.Berger, W. Bühler, R. Repges, P. Tanta) Workshop, Mainz, pp.1-15.

Elderkin, R.H., Berkowitz, D.P., Farris, F.A., Gunn, G.F., Hickernell, F.J., Kass, S.N., Mansfield, F.I., Taranto, R.G. (1977) On the steady state of an age dependant model for malaria. In : Non-linear Systems and Applications. (ed. V. Lakshmikantham). Academic Press, New York, pp. 491-512.

Feller, W. (1968) An Introduction to Probability Theory and its Applications. (3rd edn.) Vol.1, John Wiley & Sons, Inc. New York.

Fine, P.E.M. (1975) Superinfection - A problem in formulating a problem. Tropical Diseases Bulletin, 72 (b) : 475-488.

Fox, L. Mayer, D.F. (1968) Computing Methods for Scientists and Engineers. Clarendon Press, Oxford.

Freeman, R.R., Trejdosiewicz, A.J., Cross, G.A.M. (1980) Pr o tective monoclonal anti-bodies recognizing stage-specific merozoite antigens of a rodent malaria parasite. Nature. 284 : 366-368.

Cox, D. R. (1970) Renewal Theory. Butler & Tanner, London.

Garnham, P.C.C. (1966) Malaria Parasites and other Haemosporidia. Blackwell Scientific Publications, Oxford.

Garrett, Jones, C. (1964) The human blood index of malaria vector in relation to epidemiological assessments. Bulletin WHO, 30 : 241-261.

Garrett-Jones, C., Shidrawi, G.R. (1969) Malaria vectorial capacity of a population of Anopheles gambiae - an exercise in epidemiological entomology. Bulletin WHO, 40 : 531-545.

Gilles, H.M., (1981) Malaria. British Medical Journal, 283 : 1382-1385.

Ginsberg, R.B. (1971) Semi-Markov processes and mobility. Journal of Mathematical Sociology, 1 : 233-263.

Ginsberg, R.B. (1972a), Critique of probabilistic models : Application of the semi-Markov model to migration. Journal of Mathematical Sociology, 2 : 63-82.

Ginsberg, R.B. (1972b) Incorporating causal structure and exogenous information with probabilistic models : with special reference to choice, gravity, migration and Markov chains. Journal of Mathematical Sociology. 2 : 83-103.

Goodman, G.S. (1970) An intrinsic time for non-stationary finite Markov chains. Z. Wahrscheinlichkeits-theorie und Verw. Gebiete, 16 : 165-180.

Gradshteyn, I.S., Ryzhik, I.M. (1980) Tables of Integrals, Series, and Products. (4th edn.) Academic Press, London.

Gross, A.M. (1976) Confidence interval robustness with long-tailed symmetric distributions. Journal of the American Statistical Association, 71 : 409-416.

Gross, A.M. (1977) Confidence interval for bisquare regression estimates. Journal of the American Statistical Association, 72 : 341-354.

Hall, A.P (1976) The treatment of malaria. British Medical Journal, i : 323-327.

Hammersley, J.M., Handscomb, D.C. (1964) Monte Carlo Methods. Methuen, London.

Harrison, G. (1978) Mosquitoes, Malaria and Man : A History of the Hostilities since 1800. Clarke, Irwin & Co., Toronto.

Hawking, F., Wilson, M.E., Gammage, K. (1971) Evidence for cyclic development and short-lived maturity in gametocytes of P. falciparum. Transactions of the Royal Society of Tropical Medicine and Hygiene, 65 : 549-559.

Howard, R.A. (1963) Semi-Markovian Control Systems. Technical Report no.3. Research in the Control of Complex Systems. Operations Research Centre, M.I.T.

Hunter, J.J. (1969) On the moments of Markov renewal processes. Advances in Applied Probability, 1 : 188-210.

Jefferey, G.M., Eyles, D.E. (1955) Infectivity to mosquitoes of P.falciparum as related to gametocyte density and duration of infection. American Journal of Tropical Medicine and Hygiene, 4 : 781-789.

Karlin, S., McGregor, J. (1959) Coincidence probabilities. Pacific Journal of Mathematics, 9 : 1141-1164.

Karlin, S., Taylor, H.M. (1975) A First Course in Stochastic Processes. (2nd edn.) Academic Press, New York.

Kermack, W.O., McKendrick A.G. (1927) Contributions to the Mathematical theory of epidemics, I. Proceedings of the Royal Society, London, A, 115 : 700-721.

Kitchen, S.F. (1949) Falciparum Malaria. In : Malariology. (ed. M.F. Boyd), W.B. Saunders Co., Philadelphia, pp.995-1016.

Kleinrock, L. (1974) Queueing Systems. Vol.1 : Theory. John Wiley and Sons, New York.

Krafsur, E.S., Armstrong, J.C. (1978) An integrated view of entomological observations on falciparum malaria in Gambela, Western Ethiopia lowlands. Transactions of the Royal Society of Tropical Medicine and Hygiene, 72 : 348-356.

Lagakos, S.W., Sommer, C.T., Zelen, M. (1978) Semi-Markov models for partially censored data. Biometrika, 65 : 311-317.

Levy, P. (1954) Processus semi-Markoviens. Proceedings of The International Congress of Mathematics, 3 : 416-426.

Lewis, T. (1975) A model for the parasitic disease bilharzia. Advances in Applied Probabilities, 7 : 673-704.

Lotka, A.J. (1923) Contributions to the Analysis of malaria epidemiology, American Journal of Hygiene, 3 (supplement 1) : 1-121.

Lucas, A.O., Gilles, H.M. (1981) A Short Textbook of Preventive Medicine for the Tropics. Unibooks, London.

Macdonald, G. (1950a) The analysis of infection rates in diseases in which superinfection occurs. Tropical Diseases Bulletin, 47 : 907-915.

Macdonald, G. (1950b) The analysis of malaria parasite rates in infants. Tropical Diseases Bulletin, 47 : 915-938.

Macdonald, G. (1952a) The analysis of the Sporozoite rate. Tropical Diseases Bulletin, 49 : 569-586.

Macdonald, G. (1952b) The analysis of equilibrium in malaria. Tropical Diseases Bulletin, 49 : 813-829.

Macdonald, G. (1953) The analysis of malaria epidemics. Tropical Diseases Bulletin, 50 : 871-889.

Macdonald, G. (1955) The measurement of malaria transmission. Proceedings of the Royal Society of Medicine, 48 : 295-301.

Macdonald, G. (1957) The Epidemiology and Control of Malaria. Oxford University Press, London.

Macdonald, G. (1965) Eradication of malaria. Public Health Report, 80 : 870-880.

Mancini, G., Carbonara, A.O., Hermans, J.F. (1965) Immunochemical quantitation of antigens by single radial diffusion. Immunochemistry, 2 : 235-254.

Manson-Bahr, P.E.C., Apted, F.I.C. (1982) Manson's Tropical Diseases. (18th edn.) Bailliere Tindall. London.

Marcus, H.M. (1970) A stochastic model of the population dynamics of malaria parasite in the mammalian host. Ph.D. Thesis. The John Hopkins University.

Martini, E. (1921) Berechnungen und Beobachtungen zur epidemiologie und Bekampfung der Malaria. Hamburg : Gente.

McGregor, I.A., Williams, K., Voller, A., Billewicz, W.Z. (1965) Immunofluorescence and the measurement of immune response to hyperendemic malaria : Transactions of the Royal Society of Tropical Medicine and Hygiene, 59 (4) : 395-414.

McGregor, I.A. (1974) Mechanisms of acquired immunity and epidemiological patterns of antibody responses in malaria in man. Bulletin of WHO, 50 :259-266.

Mihram, G.A. (1972) Simulation, Statistical Foundations and Methodology. Academic Press, New York.

Miller, M.J. (1958) Observations on the natural history of malaria in the semi-resistant West-African. Transactions of the Royal Society of Tropical Medicine and Hygiene, 52 : 152-168.

Molineaux, L., Dietz, K. Thomas, A. (1978) Further epidemiological evaluation of a malaria model. Bulletin of WHO, 54 (4) : 565-571.

Molineaux, L., Gramiccia, G. (1980) The Garki Project, WHO, Geneva.

Moore, E.H., Pyke, R. (1968) Estimation of the transition distributions of a Markov renewal process. Annals of The Institute of Statistical Mathematics. 20 (3) : 411-424.

Moshkovskii, S.D. (1950) Basic laws of epidemiology of malaria. Moscow : AMN (In Russian).

Moshkovskii, S.D. (1967) A further contribution to the theory of malaria eradication. Bulletin of WHO 36 : 992-996.

Muir, D.A. (1975) Genetic aspects of developing resistance of malaria vector. 1. Selection pressure. Unpublished WHO document, WHO/MAL 75·856

Muir, D.A. (1977) Genetic aspects of developing resistance of malaria vector. 2. Gene flow and control pattern. Unpublished WHO document, WHO/MAL 77·885

NAG Fortran Library Manual (1977) NAG Executive, Oxford.

Najera, T.A. (1974) A critical review of the field application of a mathematical model of malaria eradication. Bulletin of WHO, 50 : 449-457.

Nardin, E.H., Nussenzweig, R.S., McGregor, I.A., Bryan, J.H. (1979) Antibodies to sporozoites; their frequent occurrence in individuals living in an area of hyperendemic malaria. Science, 206 : 597-599.

Nasell, I., Hirsch, W.M. (1973) The transmission dynamics of schistosomiasis. Communications on Pure and Applied Mathematics, 26 : 395-453.

Nasell, I. (1976a) A hybrid model of schistosomiasis with snail latency. Theoretical Population Biology, 10 (1) : 47-69.

Nasell, I. (1976b) On eradication of schistosomiasis, Theoretical Population Biology, 10(2) : 133-144.

Nasell, I. (1977) On transmission and control of schistosomiasis with comments on Macdonald's model. Theoretical Population Biology, 12(3) : 335-365.

Oyelese, Y.O. (1970) A stochastic model of epidemics involving an intermediate host (vector). Ph.D. Thesis, University of California at Berkeley.

Peter, W., Standfast, H.A. (1960) Studies on the epidemiology of malaria in New Guinea. II : holoendemic malaria, the entomological picture. Transactions of the Royal Society of Tropical Medicine and Hygiene, 54 : 249-260.

Piazza, A., Belvedere, M.G., Bernoco, D., Conighi, C., Contu, L., Curtoni, E.S., Mathing, P.L., Mayr, W., Richiardi, P., Scudeller, G., Geppellini, R. (1972) HL-A variation in four Sardinian villages under differential selective pressure by malaria. In : Histocompatibility Testing (ed.J.Dorsett and J.Colombani) : 73-84. Munksgaard : Copenhagen.

Prasittisuk, C., Busvine, J.R. (1977) DDT-resistant mosquito strains with cross-resistance to pyrothroids. Pesticide Science, 8 : 523-533.

Pull, J.H., Grab, B. (1974) A simple epidemiological model for evaluating the malaria inoculation rate and the risk of infection in infants. Bulletin of WHO, 51 : 507-516.

Pyke, R. (1961a) Markov renewal process; definitions and preliminary properties. Annals of Mathematical Statistics, 32 : 1231-1242.

Pyke, R. (1961b) Markov renewal processes with finitely many states. Annals of Mathematical Statistics, 32 : 1243-1259.

Pyke, R., Schaufele, R. (1964) Limit theorems for Markov renewal processes. Annals of Mathematical Statistics, 35 : 1746-1764.

Radcliffe, J. (1974) The periodicity of malaria. Journal of Applied Probability, 11 : 562-567.

Rao, N.R., Vig, O.P., Agarwala, S.N. (1974a) Transmission dynamics of malaria - quantitative studies 1. A stochastic model. Bulletins of the Haffkine Institute, 2 (2) : 71-78.

Rao, N.R., Vig., O.P., Agarwalla, S.N. (1974b) Transmission dynamics of malaria - quantitative studies·2. Mathematical nature of expected happenings. Bulletin of the Haffkine Institute, 2 (3) : 112-116.

Ross, R. (1911) The Prevention of Malaria (2nd edn.) Murray, London.

Ross, R. (1915) Some a priori pathometric equations. British Medical Journal, 1 : 546-547.

Ross, R. (1916) An application of the theory of probabilities to the study of a priori pathometry, I. Proceedings of the Royal Society, London, A, 92 : 204-230.

Ross, R. (1917) An application of the theory of probabilities to the study of a priori pathometry, II. Proceedings of the Royal Society, London, A, 93 : 215-225.

Russell, P.F., West, L.S., Manwell, R.D., Macdonald, G. (1963) Practical Malariology, Oxford University Press, London.

Sambasivan, G. (1975) The control of malaria. In : The Theory and Practice of Public Health (ed.W.Hobson), pp.242-256.

Segal, H.E., Wilkinson, R.N., Thiemanun, W., Gresso, W.E., Gould, D.J. (1974) Longitudinal malaria studies in rural north-east Thailand : demographic and temporal variables of infectioni Bulletin of WHO, 50 : 505-512.

Service, M.W. (1980) Guide to Medical Entomology. Macmillan, London.

Scheid, F. (1968) Theory and Problems of Numerical Analysis. McGraw-Hill, New York.

Silvey, S.D. (1975) Statistical Inference. Chapman and Hall, London.

Singer, B., Cohen, J.E. (1980) Estimating malaria incidence and recovery rates from panel surveys. Mathematical Biosciences, 49 : 273-305.

Singh, K.R.P., Brooks, G.D., Ansari, M.A.(1974) Mass rearing of mosquitoes. Journal of Communicable Diseases, 6 : 121-126.

Smaller, M.E., Sinden, R.E. (1977) Plasmodium falciparum gametocytes; their longevity and infectivity. Parasitology, 74 : 1-8.

Smith, W.L. (1955) Regenerative stochastic processes. Proceedings of the Royal Society, London, A. 232 : 6-31.

Tage, Y. (1963) On the limiting distributions of Markov renewal processes with finitely many states. Annals of the Institute of Statistical Mathematics, 15 : 1-10.

Takacs, L. (1955) Some investigations concerning recurrent stochastic processes of a certain type. Acta Math.Acad.Sci.Hungar., 6 : 101-129.

Takacs, L. (1962) Introduction to the Theory of Queues Oxford University Press, London.

Thompson, M.E. (1981) Estimation of the parameters of a semi-Markov process from censored records. Advances in Applied Probability, 13 : 804-825.

Tobie, J.E., Abele, D.C. Hill II, G.T. Contacos, P.G., Evans, C.B. (1966a) Fluorescent antibody studies on the immune response in sporozoite-induced and blood-induced vivax malaria and the relationship of antibody production to parasitamia. American journal of Tropical Medicine and Hygiene, 15 (5) : 676-683.

Tobie, J.E., Abele, D.C., Wolff, S.M., Contacos, P.G., Evans, C.B. (1966b). Serum immunoglobulin levels in human malaria and their relationship to antibody production. The Journal of Immunology, 97 (4) : 498-505.

Trager, W., Jensen, J.B. (1980) Cultivation of erythrocytic and exoerythrocytic stages of plasmodia. In : Malaria (ed.J.P.Kreier) 2 : 271. Academic Press, London.

Waite, H. (1910) Mosquitoes and malaria. A study in the relation between the number of mosquitoes in a locality and the malaria rate. Biometrika, 7 : 421-436.

Weiss, G.H., Zelen, M (1965) A semi-Markov model for clinical trials. Journal of Applied Probability, 2 : 269-285.

Willis, D.M. (1964) The statistics of a particular non-homogeneous Poisson process. Biometrika, 51 : 399-404.

World Health Organization (1981) Weekly Epidemiological Record, No.10.

Yekutiel, P. (1960) Problems of epidemiology in malaria eradication. Bulletin of WHO, 22 : 669-683.

Yoshida, N., Nussenzweig, R.S., Potocnjak, P., Nussenzweig, V., Aikawa, M. (1980) Hybridoma produces protective antibodies directed against the sporozoite stage of malaria parasite. Science, 207 : 71-73.