Assessment of seasonal and climatic effects on the incidence and species composition of malaria by using GIS methods

Thesis submitted for the degree of Doctor of Philosophy

Ali-Akbar Haghdoost

London School of Hygiene and Tropical Medicine

Faculty of Medicine

University of London

April 2004

Abstract

Malaria is highly dependent on climate and environmental factors. This thesis incorporates environmental and climatic factors into mathematical and geographic information system (GIS) models in order to assess the feasibility of an early warning system in a strongly seasonally transmitted region in Iran. It also measures *Plasmodium* spp interactions through meta-analysis, modelling, and further analysis of a large epidemiological dataset.

The first part of the thesis assesses the feasibility of malaria prediction models based on ground climate and remote sensing data. Predicted values were typically extrapolated from the previous month's data; adding ground climate data can improve these predictions by around ten percent. Predictive variables for these models are readily available in the field, so an improvement of even a few percent makes them feasible. However, more ground climate data are needed for prediction at finer than district spatial scales.

The second part of the thesis measures interactions between malaria species. A systematic literature review and meta-analysis assessed the heterogeneity of interaction terms between malaria species. Mathematical models assessed the effects of within-population heterogeneity in infection risks. Finally, data from a large epidemiological study in a highly malaria-endemic area (Garki, West Africa) were analysed cross-sectionally and longitudinally.

Random-effect meta-analysis produced a summary OR between *P. falciparum* and *P. vivax* of less than one (0.6, 95% CI 0.46-0.8). The very

wide range of ORs seen between studies (0.02 to 10.9) could be explained partly by species prevalence and the temporal span of studies.

Mathematical models indicated that within-population heterogeneity in infection risks may, by itself, explain ORs as great as ten or more.

Longitudinal analysis of the Garki data produced lower ORs than those from cross-sectional analysis. *P. falciparum* had suppressive effects on the other species. In addition, *Plasmodium* spp interactions highly depend on subject age and the temporal and spatial distribution of species. In conclusion, heterogeneity in infection risks, due to heterogeneity either in acquired immunity or in exposure risk, is the most important factor on interactions between *Plasmodium* spp.

Finally, it seems that species-specific models would improve the predictions due to the different impacts of climate on the transmission of species and the interaction between them.

Contents

PART TWO
INTERACTION BETWEEN *PLASMODIUM* SPP

Dedication

I whole-heartedly dedicate this thesis expressing affection and gratitude to my wife, Atoosa and my daughter, Bahar for their understanding.

Acknowledgements

I should morally dedicate my grateful acknowledgements to Dr Neal Alexander for his continuous help and valuable advice. Also I would like to thank Dr Jonathan Cox who scientifically supported me particularly in GIS and RS issues; Dr Tom Smith who facilitated my access to the Garki Project data and gave me helpful advice especially about the interaction between malaria species; Dr Shakoor Hajat for his useful guidance on time series.

I would like also to convey sincere thanks to Dr Iraj Sharifi, Kambiz Bahadini, and Jelaladin Mahzooni for their kind and effective co-operation on field data collection, and all other staff in Kerman Medical University who help me in this project particularly F. Kazemi, E. Bamorovat, T. Shaban-Nejad, M. Safari and M. Mashayekhi.

The GIS part of this project is financially supported by DFID Malaria Programme (London School of Hygiene and Tropical Medicine) and the field data are being collected by the contribution of Kerman Medical University (Iran).

Abbreviations

ABER	Annual Blood Examination Rate
AFI	Annual *falciparum* Index
An.	*Anopheles*
ANOVA	Analysis of Variance
APC	Antigen Presenting Cell
API	Annual Parasite Index
ASO	Annual Sum of Over-estimation
ASU	Annual Sum of Under-estimation
AVI	Annual *vivax* Index
CCD	Cold Cloud Duration
DA	Drug Administration
DME	Digital Elevation Model
DNA	Deoxyribonucleic Acid
EMRO	Eastern Mediterranean Region Organisation
FP	Fractional polynomial
GIS	Geographical Information System
GP	General Practitioner
GPS	Global Positioning System
HW	Health Worker
IgG	Immunoglobulin G
IgM	Immunoglobulin M
kg	kilogram
km	kilometre
LR test	Likelihood Ratio test
LST	Land Surface Temperature
m	metre
MDA	Mass Drug Administration
mg	milligram
MHC	Major Histocompatibility Complex
mm	millimetre
NDVI	Normalised Differentiate Vegetation Index
NPV	Negative Predictive Value
OR	Odds Ratio
P. f.	*Plasmodium falciparum*
P. m.	*Plasmodium malariae*
P. o.	*Plasmodium ovale*
P. v.	*Plasmodium vivax*
PCR	Polymerase Chain Reaction
PHC	Paramount Health Centre

PPV	Positive Predictive Value
RHC	Rural Health Centre
RIGLS	Restricted Iterative Generalised Least Squares
ROC	Receiver Operating Characteristic
RR	Risk Ratio
RS	Remote Sensing
RSO	Relative Sum of Over-estimation
RSU	Relative Sum of Under-estimation
SAR	Sum of the Absolute Residuals
SD	Standard Deviation
SE	Standard Error
Sen	Sensitivity
SO	Sum of Over-estimation
Spe	Specificity
SPR	Smear Positive Rate
SQL	Structure Query Language
SSD	Sub Sub District
SU	Sum of Under-estimation
WHO	World Health Organisation
WMH	World Meteorological Organisation
µl	Microlitre

List of tables

List of figures

CHAPTER 1

1. General introduction

1.1. Malaria

Malaria is a parasitic infection transmitted to humans through the bites of infected female *Anopheles* mosquitoes. The resulting disease in humans can be devastating. After travelling rapidly through the bloodstream to the liver, the parasite emerges again into the blood stream, finally to settle in the red blood cells, where it multiplies and emerges in bursts of new organisms. These parasites, because of their large numbers, can cause particular damage to the nervous system, liver, and kidney *[1]*.

Malaria occurs in over 100 countries and territories. More than 40% of the people in the world are at risk. Large areas of Central and South America, Africa, the Indian subcontinent, Southeast Asia, the Middle East, and Oceania are considered malaria-risk areas *[2]*.

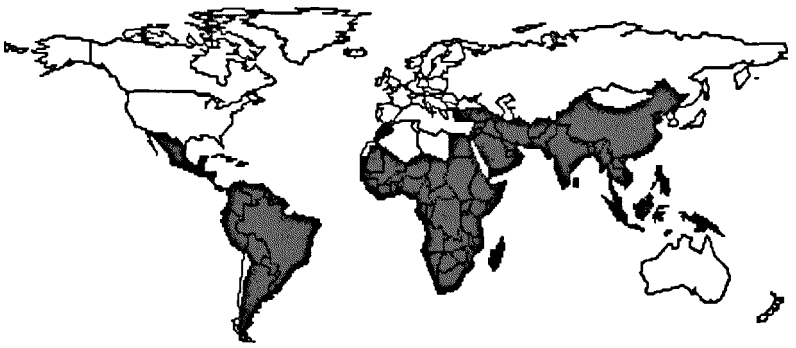

Figure 1-1: The distribution of malaria in the world *[WHO/TDR 2003]*

The World Health Organization (WHO) estimates that yearly 300-500 million cases of malaria occur and more than one million people die of malaria. In young children and adults who have not recently been infected (and therefore have not developed natural immunity), this cycle can result in death within hours from cerebral malaria. Others die later from overwhelming anaemia or liver and kidney failure. Untreated, up to 20% of persons infected with *falciparum* malaria will die *[WHO/TDR 2003]*.

There are four species that can infect humans: *P. falciparum, P. vivax, P. ovale,* and *P. malariae.* They have a life cycle which is split between a human host and an insect vector. The *Plasmodium* spp, with the exception of *P. malariae* (which may affect the higher primates) are exclusively parasites of man *[1]*.

The mosquito is always the vector, and is always an *Anopheline* mosquito, although, out of the 380 species of *Anopheline* mosquito, only 60 can transmit malaria. Only female mosquitoes are involved as the males do not feed on blood *[2]*.

Malaria has a complex life cycle. Infected female mosquitoes inject malaria sporozoites when they bite, and the sporozoites are carried to the liver where they rapidly infect liver cells. Without causing symptoms, these sporozoites undergo a radical change and multiply furiously for the next 4-5 days. Tens of thousands of asexual stage merozoites are released from each infected liver cell, each of which rapidly targets and invades a red blood cell. Every few days, the merozoites multiply ten-fold and burst out to infect other red blood cells. This cyclic and massive increase in parasite burden gives rise to the clinical disease we recognise as malaria *[1]*.

In the absence of immunity or drug treatment, death can occur within hours of noticeable symptoms. If death does not occur and infection continues, some of the parasites further differentiate into a form that is infectious for mosquitoes (gametocytes), thus permitting the life cycle to continue *[2]*.

The diagnosis of malaria is confirmed by blood tests and can be divided into microscopic and non-microscopic tests. Microscopic tests involve staining and direct visualisation of the parasite under the microscope. Non-microscopic tests involve identification of the parasitic antigen or the anti-plasmodial antibodies or the parasitic metabolic products *[2]*.

Many different approaches have been taken to prevent the cycle of disease and mortality from malaria. These approaches can be classified as follow *[2]*:

1. Control of mosquitoes, e.g. by eliminating stagnant water where they can breed, or by using pesticides
2. Limiting human exposure to the infected mosquito
3. Prevention of the disease through prophylactic use of anti-malarial drugs

Nonetheless, despite of a great deal of research, malaria still is one of the main international health problems s and there are many questions without any precise responses, such as questions about the best controlling methods, immunity against malaria, and prediction of malaria based on climate and remote sensing data.

1.2. Overview of thesis

This thesis explores two related topics. The first topic is the effects of climate on malaria. The second topic is the interaction between *Plasmodium* spp. These are related because the interaction between species may affect temporal and spatial variations of species. In addition, the discrepancies between species distributions due to the climate effect may influence their interactions.

The number of malaria cases is modelled based on meteorological and remote sensing data in Part One. This part explains the effects of climate on malaria and models species-specific incidence in an endemic area of Iran. Furthermore, the differences between the spatial and temporal distributions of *P. falciparum* and *P. vivax* are explored.

Part Two explores the interaction between *Plasmodium* spp and has four chapters. The first chapter (number 4) is an introduction and explanation of the possible sources of interactions between species. The second chapter assesses the interactions between *P. falciparum* and *P. vivax* in a systematic review of the literature and explores the sources of heterogeneity in this interaction. The third chapter models the association between species based on the variation of infection risks within a population. The fourth chapter assesses the interaction between *Plasmodium* spp in the Garki Project, a very large epidemiological study of malaria in Africa, with both cross sectional and longitudinal data.

The last chapter of this thesis discusses the main findings and links the results. It also explores the study's limitations and suggests topics for further studies.

1.3. Rationale

A precise risk map can improve the effectiveness of control programmes. In addition, malaria transmission is highly affected by climate. Therefore, it seems that predictive models based on meteorological factors may improve the effectiveness of control programmes. There is a great deal of research on this topic, but most of it has explored this association in epidemic areas, mainly in Africa. Therefore, the feasibility of an early warning system based on ground meteorological and remote sensing data is assessed in a district in Iran with seasonal malaria (Part One).

The interaction between *Plasmodium* spp has been studied at different levels, from molecular up to human populations. Nonetheless, there is not any

systematic literature review or meta-analysis on this topic. Also, most epidemiological studies have analysed this topic cross-sectionally. The second part of this project assesses some of these aspects in the interaction between *Plasmodium* spp longitudinally as well as cross-sectionally. In addition, it presents the results of mathematical models.

Considering these two parts together, it may be possible to explain some of the apparent interaction between species by spatial and climatic variabilities. Moreover, the observed discrepancies between temporal variations of *Plasmodium* spp may be explained either by their different sensitivities to climate or the interaction between species; i.e., the suppressive effect of one species on the others.

Part One: Modelling of malaria based on climate and remote sensing data

Prospectus

This part addresses the association between climate and malaria risk and has two chapters. The first chapter reviews the literature and discusses the possible pathways through which climate may affect malaria risk. It also discusses the application of geographical information system and remote sensing to public health.

The second chapter presents the results of statistical modelling to which assess the feasibility of an early warning system in Kahnooj, a district in Iran with seasonal malaria. It describes the health system in Kahnooj and presents the malaria situation. Then, it assesses the feasibility of an early warning system in predicting malaria using meteorological variables and remote sensing data.

CHAPTER 2

2. Climate effects on malaria

2.1. Introduction

Seasonal and climatic variation affects the occurrence of epidemics of malaria, as well as long-term trends. High temperature, humidity and rainfall are mostly considered as the main risk factors for malaria outbreaks in epidemic areas *[2]*. It is well established that climate is an important determinant of the spatial and temporal distribution of vectors and pathogens *[3]*. Ndiaye et al. (2001) showed the relation between climate variability, both seasonally and inter-annually in Senegal, and the variability in the number of deaths attributable to malaria *[4]*.

The approximate average monthly temperature cut off point between an epidemic and a non-malaria zone in Africa is indeed around 18°C; and 22°C allows stable transmission *[5]*. The minimum temperature is also important in the transmission of malaria. The minimum temperature iso-line is around 4°C and for stable transmission 6°C *[5]*. In addition, other factors such as wind and the duration of daylight can also be significant *[6]*.

Furthermore, the duration of the warm season is also important. It seems that unusually high temperatures at the end of the normal malaria season

prolong transmission and substantially increase the number of malaria cases (particularly those infected with *P. falciparum*) *[7]*.

2.2. Direct effects of climate

Mosquito abundance depends on the rate at which insects are produced from their breeding sites, and their survival rates. Higher temperatures speed up the development of adult mosquitoes, which live longest between 25°C and 35°C. At very low and very high temperatures, mosquitoes have shorter lives *[2]*.

The person-biting rate is a measure of the number of times that each person gets bitten each day. This number is dependent on the frequency of mosquitoes per person, the feeding behaviour of the mosquitoes, and human behaviour.

Climatic conditions and temperature in particular, directly influence mosquito development, feeding frequency and longevity, as well as the time in which the parasite develops inside the mosquito. Other environmental factors such as vegetation and breeding sites are indirectly influenced by climate conditions *[8]*.

Vectorial capacity describes the intensity of malaria transmission *[9]* and is defined as the mean number of potentially infective bites that will be delivered by the vectors feeding on a single infectious host in one day. Vectorial capacity (C) is a function of four components *[10]*:

$$C = \frac{ma^2 p^n}{-\ln(p)}$$ Equation 2-1

where m is the frequency of female mosquitoes per person; a is the frequency of blood feeding by one mosquito in one day; n is the duration of sporogonic phase in days; and p is the probability of a vector surviving one day.

Suitable climate, particularly high temperature, speeds up the development of *plasmodia* parasites within mosquitoes. However, it should be borne in mind that without sufficient breeding sites and effective contact between mosquitoes and humans, higher temperature will not matter *[11]*.

The critical density of mosquitoes for transmission of malaria can be expressed as *[8]*:

$$m^* = k\left[\frac{-\log(p)}{a^2 p^n}\right] \qquad \text{Equation 2-2}$$

where m^* is the critical number of female mosquitoes per human; and k is a constant which is dependent on the recovery rate of human and efficiency of transmission between human and mosquitoes.

All parameters in Equation 2-2 depend directly on climate factors except k, which might be affected indirectly by environmental factors.

The pathways by which climate affects malaria incidence are described in the following sections.

2.2.1. Temperature

Temperature is important because it governs *[12]*:

1. The rate at which mosquitoes develop into adults,

2. How frequently they blood feed (and, therefore, acquire parasites),
3. The survival rate of adult mosquitoes,
4. The incubation time of parasites in the mosquito,

1. The rate at which mosquitoes develop into adults: as shown in Figure 2-2, a small increase in temperature at low temperatures has a greater effect on the development time than similar changes at higher temperatures. In other words, the downward slope is steeper at lower temperatures. Mosquito populations generally increase more rapidly in warmer climates. At extremely high temperature, over 30°C, development actually slows down *[13]*. However, the availability of breeding sites is the main limiting factor of the abundance of mosquitoes *[14]*. For example, in spite of the cold climate of northern Canada, the number of mosquitoes in some places is surprisingly high due to availability of breeding sites *[15]*.

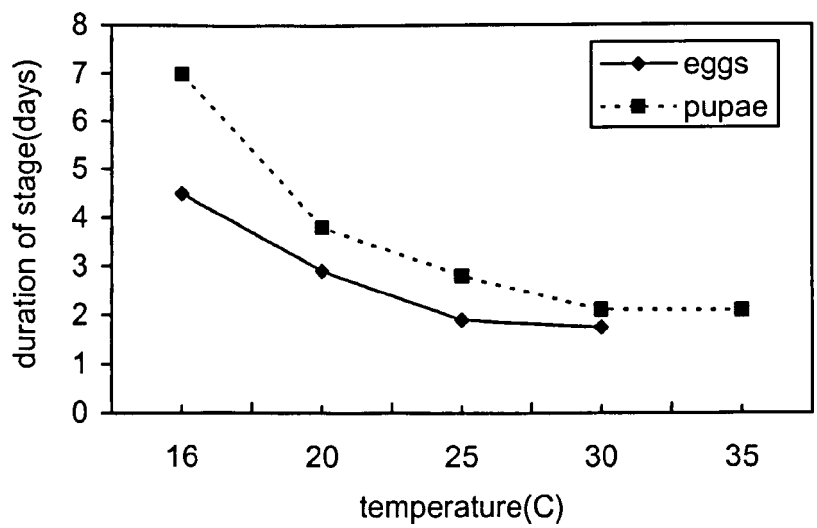

Figure 2-1: Development of eggs and pupae of *An. minimus* at different temperatures *[16]*

2. Frequency of blood feeding: most *Anopheles* are "anautogenous", requiring at least one blood meal to produce each clutch of eggs, therefore the search for blood is an ongoing, repetitive process *[17]*. Each cycle starts with an unfed adult, passing through a blood-fed, half gravid and gravid condition; then after oviposition the female is again unfed and seeks another blood-meal (gonotrophic cycle). In an idealised cycle, feeding and egg-laying occur once per cycle, irrespective of its duration *[14]*. During the dry-hot season the act of oviposition, and probably the maturing of the eggs, are delayed, and more than one blood meal is taken per gonotrophic cycle *[18]*. The feeding frequency and the length of the blood-digestion phase are dependent on the ambient temperature *[14]*. However, endophilic vectors, resting in more sheltered environments, are less sensitive to ambient temperature *[15]*.

In the wild, tropical mosquitoes usually bite at regular intervals of 2-5 days. But small changes in temperature have substantial effects *[14]*. For example, Lindsay (1996) showed that *An. gambiae* and *An. funestus* feed every two days at around 25°C, but only every three days at lower temperatures *[14]*.

In general, the length of gonotrophic cycle (n_u days for the uth cycle) and temperature can be related as follows:

$$n_u = \frac{f_u}{T - g_u} \qquad \text{Equation 2-3}$$

where f_u is a thermal sum, measured in degree-days, representing the accumulation of temperature units over time to complete the development, g_u a threshold below which the development ceases, and T is the ambient temperature. Table 1 shows these parameters for the first gonotrophic cycle

(u=1) *[14]*. Figure 2-3 shows the estimated duration of the gonotrophic cycle of the three main malaria vectors in the research setting of Kahnooj, Iran, according to Equation 3 and Table 2-1.

Table 2-1: Threshold temperatures and duration of the first gonotrophic cycle (u=1) for the three main malaria vectors in Kahnooj.

Parameter	*Anopheles* species		
	maculipennis[1]	*culicifacies*[2]	*stephensi*[2]
f_1 (degree days)	36.5	29.7	43.4
Threshold temperature(°C)	9.9	12.6	8.9

1 Data from Detinova (1962) *[19]*
2 Data from Mahmood and Resien (1981) *[20]*

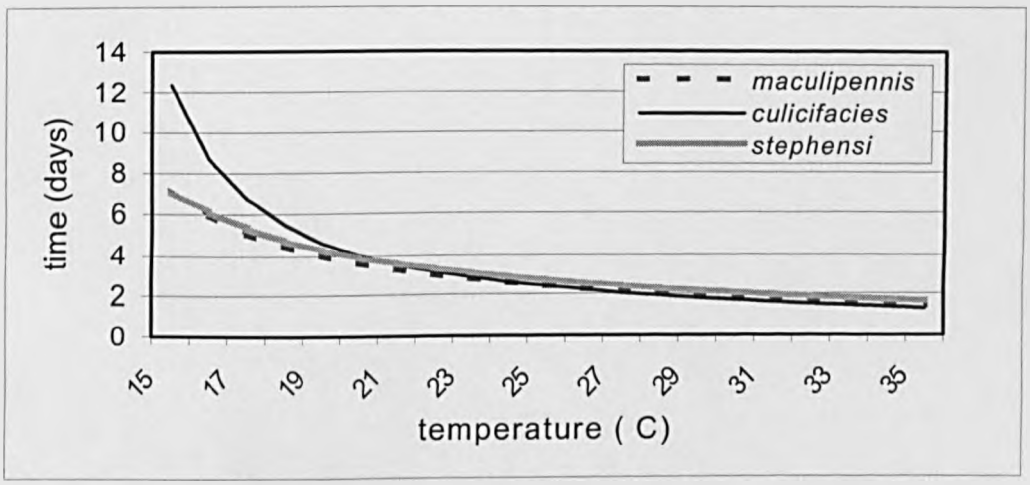

Figure 2-2: The duration of the first gonotrophic cycle of three *Anopheles* species as main malaria vectors in Kahnooj, Iran

3. Adult mosquito survival: experimental studies show that mosquitoes can survive many months in cold climates *[6]*, while there may be a negative effect on survival of adult mosquitoes at high temperatures *[21]*. Experimental studies show that the thermal death of mosquitoes occurs around 40-42°C *[22,23]*. In the field, *Anopheles* have the highest survival at 25-35°C *[21]*.

Assuming constant humidity, the probability of a mosquito surviving one day can be estimated as follows *[5]:*

$$p = e^{-1/(-4.4+1.31T-0.03T^2)}$$ Equation 2-4

4. The incubation time of parasites in the mosquito: the length of time required for the development of the malaria parasite within a mosquito (n_t) is also linked to ambient temperature, and can be summarised thus *[19]:*

$$n_t = \frac{M}{(t-m)}$$ Equation 2-5

where n_t is the length of sporogonic cycle in days, M is a constant (degree days representing the accumulation of temperature units over time, 105 for *P. vivax* and 111 for *P. falciparum*). In this equation t represents the mean monthly temperature, and m the lowest temperature at which the sporogonic cycle can be completed for the species of malaria (14.5°C for *P. vivax* and 16°C for *P. falciparum*) *[19,24].*

The minimum temperatures for the development of *P. falciparum* and *P. vivax* are about 18 and 15°C respectively *[25].*

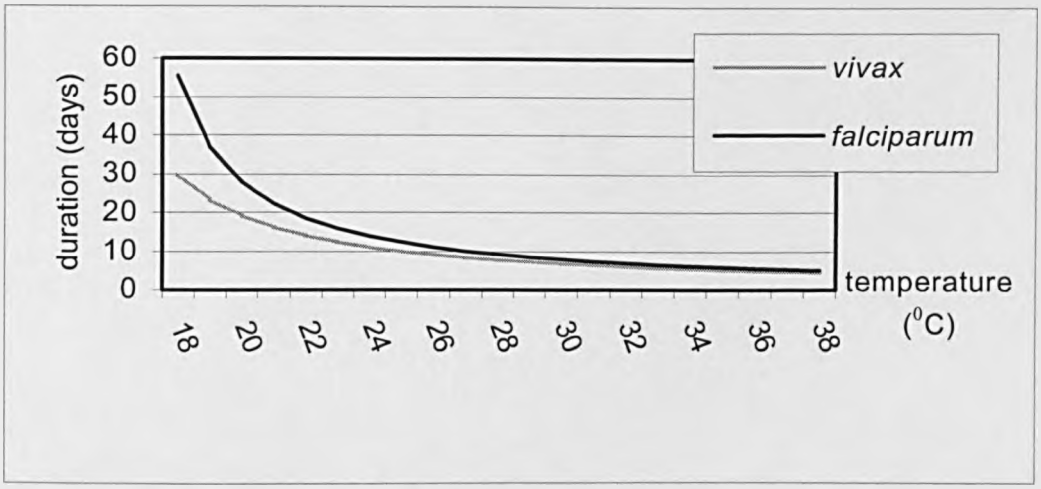

Figure 2-3: Duration of sporogonic cycle (the period required for parasites to become infective in the vector) of *P. vivax* and *P. falciparum* at different ambient temperatures (Equation 2-5)

2.2.2. Humidity

Humidity is one of the factors which has a direct effect on the survival of mosquitoes. Survival rate might be reduced when hot weather is accompanied by low humidity. But in areas with this type of climate, such as semi-arid parts of Sudan, local species have adapted themselves *[6]*.

Humidity also affects the risk of exposure to vectors (Section 2.3.3).

2.2.3. Precipitation

Rain provides the breeding sites for malaria mosquitoes and helps create a humid environment, which prolongs the life of vectors *[12]*. Craig (1999) shows that at least 80mm per month precipitation for five months is needed for stable malaria transmission in Africa *[5]*. Ndiaye et al. (2001) showed that the correlation between the variability in August rainfall and the

variability in the number of deaths attributed to malaria between August and December was strong, positive and statistically significant (r=+0.61, p=0.02). In addition, highly significant cross-correlations were found between monthly rainfall series and monthly mortality series at one- and two-month lags (r=+0.43, p=0.0004 for one-month lag; r=+0.26, p=0.03 for two-month lag) *[4]*.

On the other hand, heavy rains may have a flushing effect, cleansing breeding sites of their mosquitoes *[5,6]*. Heavy autumnal rain fall in Algeria, for example, normally brings the malaria transmission season to an end by flushing out the breeding sites of *An. labranchiae [2]*.

2.3. Other considerations related to climate

2.3.1. Deforestation

Climate changes may lead to adaptation of in local human behaviour. Forest clearance and agricultural activity provide favourable conditions for those insects that prefer irrigation ditches, wells and temporary ground pools exposed to full sunlight as breeding sites, such as many types of *Anopheles [26]*. Agricultural fertilisers can promote the growth of algae and other larval nutrients, although herbicides may eliminate them altogether *[6]*.

2.3.2. Migration and urbanisation

Drought, flooding or economic factors can cause mass population movements *[26]*. Infected people can introduce malaria to non-endemic areas. Rapid population expansion can cause breakdown in public health services. In addition, extensive water storage and inadequate water disposal can lead to disastrous surges in the number of malaria mosquitoes. Furthermore in large cities and camps, zoophilic species might be encouraged to feed on people due to dense human population and the absence of cattle *[6]*.

2.3.3. Changing human behaviour

The lifestyle of people is dependent on the climate. Usually people wear less clothes in warm and humid climates, and prefer to work and rest in open areas. Therefore, they are more exposed to mosquito bites. On the other hand, due to the abundance of insects they might use more bed-nets or other protective methods, which decrease the risk of effective exposure *[26,27]*.

2.3.4. Natural disaster and conflict

Natural disasters such as drought and flood might disrupt the health infrastructures and change human life, and so might create an optimum condition of any types of epidemics *[28-34]*.

Flooding often causes disruption of breeding sites and temporary reduction of vectors. But it never eliminates the vectors, so that high rainfall is still considered optimal for transmission *[5,26]*.

2.4. Prediction of malaria by meteorological factors

2.4.1. Basic concepts

Understanding the linkages between meteorological and ecological changes, as determinants of disease, emergence and redistribution will ultimately help to optimise preventive strategies *[35]*. Craig et al. (1999) demonstrated that a simple climate-based model can be used to define the crude distribution of malaria transmission in Africa *[5]*. However, the long term variation in malaria risk based on climate changes must be interpreted based on local environmental conditions, socio-economic development, and malaria control programmes or capabilities *[8]*. Reiter (2001) emphasised that human activity and general social behaviour can be much more important than the effect of climate on malaria incidence. Therefore, it is inappropriate to use only climate-based models to fully predict future prevalence *[6]*.

Also, due to the complexity in the relationship between the variables, looking at only one meteorological factor and ignoring the effect of others might produce distorted results. Malaria in highly endemic regions is less sensitive to climate changes. However, a small increase in temperature in highly seasonal areas can disproportionately enhance malaria transmission

[8,14]. A study in Mali, which has high temperatures, showed that a rainy season of only three months is enough to sustain transmission of *P. falciparum*. In contrast, in southern and eastern parts of Africa with cooler weather, at least five consecutive months of rain are needed to enable mosquitoes to achieve the abundance required to sustain malaria *[5]*. In western Kenya, malaria outbreaks have occurred at altitude of 2000m when the mean monthly temperature exceeded 18°C and rainfall reached more than 15 cm per month *[36]*.

The results of another study in Pakistan showed that *P. falciparum* is more susceptible to climate changes than *P. vivax [37]*. All these examples support the idea that a comprehensive view of the local epidemiology of malaria is needed to make an efficient model based on the meteorological factors.

Microclimate is very important in malaria transmission *[38]*. The mean daily ground temperature in dense forest is around 10°C less than adjacent open areas. In addition, indoor and outdoor temperature can be several degrees higher or lower, depending on the season, house design and construction material. Mosquitoes use a variety of strategies to exploit the timing and location to maximum advantage *[6]*.

Distribution maps of vector types are important for understanding the epidemiology of the diseases they transmit. Bayoh et al. (2001) developed a model to predict the chromosomal forms of *An. gambiae*, the principal vector of malaria in West Africa, based on climate data, with more than 80% accuracy. Modelling demonstrated that climate affects not only the distribution of vector species but also their genotypes *[39]*.

In Africa, most malaria deaths occur at the end of the rainy season. Greenwood and Pickering (1993) explained this finding by the effect of climate on the malaria species, immunity of the population against malaria, drug resistance and other socio-economical factors *[40]*.

Changes in precipitation and temperature can have a marked effect on the intensity of transmission *[14]*. Multivariable analysis of the malaria in a district in North West Pakistan showed a significant relationship between the annual frequency of *P. falciparum* cases and rainfall in September and November, temperature in November and December, and humidity in December *[37]*.

2.4.2. Limitation of modelling based on climate

Human vulnerability is the product of immunity, poverty and behaviour *[14]*. Economic, social and political factors are very important in the frequency of malaria cases, but are outside the scope of nearly all statistical and mathematical models. For example, Thomson et al. (1994) showed that in The Gambia the most malarious areas were those with fewest mosquitoes. This paradox was explained by the habits of people; in general, people slept under bed-nets only in villages with high densities of mosquitoes *[27]*.

In addition, there are some assumptions in the mathematical models which might not be true in practice. For instance, the risk of contact between people and vectors is usually positively skewed, i.e., a small proportion of people have a much higher contact rate than the rest of population. Taking

account of this heterogeneity usually gives a different result, but limitations of data often prevent it being included.

Furthermore, there are some uncertainties in measuring model parameters, especially with remote sensing (RS) data *[41]*. These types of errors may significantly reduce the accuracy of model results.

Nevertheless, it should be mentioned that from a practical point of view, a simpler model based on existing field data is usually preferred to an expensive and complex model with more accuracy.

2.5. Geographical Information System (GIS)

2.5.1. Definition

GIS is a computer-based method for storing, mapping, visually analysing and reporting data in spatial as well as temporal formats *[42]*.

In relation to vector-borne diseases, GIS enables us to identify and analyse factors that may explain part of the temporal and spatial distribution of vectors as well as the disease. In this way, climatic variables and land cover are the most important explanatory variables, and are being increasingly applied to the study of vector-borne diseases *[41,42]*.

2.5.2. Remote sensing (RS)

Remote sensing methods enable us to assess some property of an object without direct contact with it.

Nowadays, satellites are scanning the earth surface and capturing environmental data at a range of resolutions (spatial, temporal, spectral and radiometric) *[42]*. The satellite sensors measure radiation in various regions of the electromagnetic spectrum emitted, reflected or scattered from each point of the earth. Due to differences in the optical behaviour of different objects, it is possible to characterise them through their spectral signatures. In passive RS, the sensor measures the solar radiation reflected by the Earth surface, or radiation emitted by the Earth itself. In active RS the sensor detects of its own reflection of the generated long wavelength signal (3-30m) by land surface features *[43]*.

In relation to vector-borne diseases, the main environmental proxies derived from meteorological satellite data are: Cold Cloud Duration (CCD) an estimator of rainfall; Normalised Difference Vegetation Index (NDVI), an indicator of land-cover and water tables; altitude and Land Surface Temperature (LST) *[42]*.

2.5.3. Application of GIS in malaria

Sipe and Dale (2003) *[44]* reviewed the GIS and malaria literature and divided the publications into the five categories outlined below:

1. Mapping malaria incidence/prevalence
2. Mapping the relationships between malaria incidence/prevalence and other potential related variables
3. Using innovative methods of collecting data such as remote sensing (e.g., GIS)

4. Modelling malaria risks
5. General commentary and reviews of GIS used in malaria control and research

Studies can be classified into two main groups according to types of outcome variables. Due to the direct effect of environmental factors on malaria vectors, the first group tries to model the abundance of vectors. The second group explains the effect of environmental factors on the frequency of malaria cases.

Klinkenberg and Van Der Hoek (2003) *[45]* and Noor et al. (2003) *[46]* used GIS methods to predict the malaria risk. Hay et al. (2000) reviewed the current status of GIS and RS as new tools to improve programmes for the control of malaria and its vector in sub-Saharan Africa *[47]*. He used remote sensing data to predict the population dynamics of arthropod vectors *[48]*. In a study in Mexico, the prevalence of *Anopheles* was predicted by RS data using two statistical approaches, discrimination analysis and regression. Both showed 70% accuracy in prediction of high prevalence villages *[49]*.

Despite some limitations of these techniques, Sharma et al. (1996 and 1997) showed in his papers that GIS and RS are very useful in mapping the major breeding sites, recording temporal changes and estimating larval production in a cost-effective and timely manner *[50,51]*.

Roberts et al. (1994 and 1996) *[52,53]*, Rejmankova et al. (1995) *[54]* and Lindsay et al. (1998) *[55]* used RS data to predict the malaria vector distribution by different approaches.

Malaria transmission in the Red River basin in China is primarily determined by migration of people, environmental variables, particularly

altitude, paddy and forest *[56]*. Connor et al. (1998 and 1999) used satellite data to generate a risk map of malaria and develop an early warning system *[57,58]*. In another study in 26 villages in The Gambia, clinical and entomological data were linked to RS data. Malaria risk and the natural immunity against malaria were found to be closely related to the distance between villages and breeding sites *[59]*. A study in Kenya showed a strong association between the number of paediatric severe malaria cases and NDVI (r^2=0.71) *[60]*. Hay et al. (1996) compared RS data with ground data and discussed the potential source of errors. He also showed in another review that using RS methods improves the efficacy of malaria control programmes *[43]*. All these studies showed the application of RS data on the prediction of malaria.

It is expected that vector abundance and the frequency of malaria cases are strongly correlated. On this assumption, mosquito abundance should be a good proxy of malaria burden. However, some studies do not support this assumption. There are some malaria free areas with uninfected vectors in Europe, North America and Australia *[43,53,58]*. Human behaviour and other socio-economic factors are also important in determining the incidence of malaria (Section 2.3).

Due to this discrepancy between abundance of vectors and frequency of malaria cases, from the public health policy point of view, the modelling of malaria cases might be more reliable. However, due to greater validity and availability, most of studies modelled the abundance of vectors. Nonetheless, the following chapter models the incidence of malaria to assess the validity of predictions based on RS and ground climate data.

Linear regression, logistic regression and discriminant analysis are the main statistical methods which have been used to assess the relationship between RS data and the frequency of vectors or cases. However, special methods are needed to take into account the dependencies caused by auto-correlation between data consecutive in time or space *[61]*. In addition, new methods such as neural networks and fuzzy rule-based systems are being developed, although their applications are currently limited *[42]*. In the next chapter, the Poisson regression method (adjusted for dependencies) was used to predict the number of malaria cases.

CHAPTER 3

3. Feasibility of an early warning system

3.1. Overview

This chapter uses ground and remote sensing data to model temporal and spatial malaria risks in Kahnooj, a malaria endemic area in southern Iran, between 1994 and 2002.

The main objective is to assess the feasibility of an early warning system based on meteorological and remote sensing data for predicting malaria cases in an area of highly seasonal transmission.

Malaria data were extracted from the surveillance system, which records the date, location and the species of infection for each malaria case. In addition, the type of surveillance (active or passive), nationality and the names of cases were recorded.

The meteorological data were obtained from the synoptic centre in Kahnooj city, and remote sensing images. The ground data contain the daily minimum and maximum temperature, mean daily relative humidity and rainfall. The remote sensing data contain the mean land surface temperature and vegetation index over 10 or 30 days; with two spatial resolutions: 30x30m, and 8x8km.

Using Poisson regression models, the numbers of expected cases were predicted based on temperature (ground and remote sensing), humidity, rainfall, vegetation index, seasonality, time trend and auto-correlation between temporal risks.

The data were grouped by month or ten day period; then the aggregated records were randomly allocated to modelling (75%) and checking (25%) (Section 3.4.8). The parameters were estimated based on the modelling data. The accuracy of models was checked by comparing the fitted and observed values in the checking data (Section 3.4.8).

This chapter also assesses the risk factors for local transmission. A map of local transmission is a useful tool for the health system to identify and concentrate its activities in high risk areas to prevent outbreaks. The local transmission in each village was defined as the presence of at least two conspecific malaria cases in one month, or one conspecific case in each of two consecutive months. The sensitivity and specificity of models in detection of locally transmission were computed using the ROC method.

In addition to sections which address the main objective, the relapse rate of *P. vivax*, and treatment failure risk of *P. falciparum* are presented in Section 3.7.3.

3.2. Research setting

The following sections describe the geography, including climate, of Iran and the study area. They also describe the health system and malaria surveillance.

3.2.1. Geographical description of Iran

Iran is located in the Middle East with an area of 1.6 million square kilometres. It extends north to Azerbaijan, Turkmenistan, Armenia and the Caspian Sea; east to Afghanistan and Pakistan; south to the Persian Gulf and the Oman Sea; and west to Turkey and Iraq.

Mountain ranges divide Iran into four different climatic and biotic regions:

1. Caspian Sea littoral between the northern slope of the Alborz mountains and the Caspian Sea. This is a narrow strip of forested land, with a Mediterranean climate (10-35°C and 70%-100% relative humidity)
2. The central plateau extends to the east and south-east with hot and dry climate in summer and cold, snow-bound winters (0-40°C and 0-40% relative humidity)
3. Persian Gulf and Khuzestan plain in the south and south-west has a tropical climate (15-40°C and 40%-80% relative humidity)
4. Mountainous area in north-west area with very cold winters (-10 to 25°C and 30%-60% relative humidity)

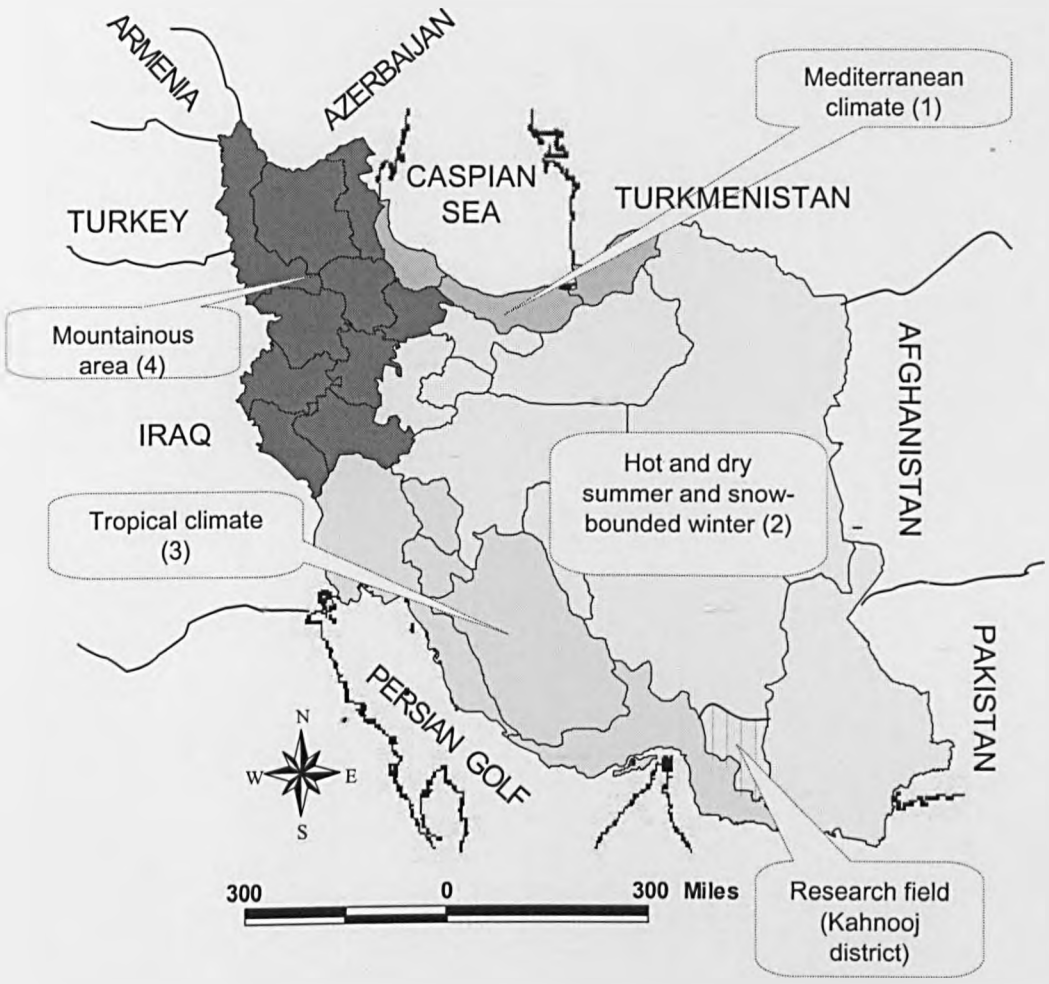

Figure 3-1: Map of Iran, showing different climatic and biotic regions. Kahnocj is shown in the south-east.

3.2.2. Description of Kahnooj

3.2.2.1. Climate

Kahnooj is a district of Kerman province in southern Iran with hot dry weather and an area of 32,000km^2. Three seasons may be recognised: wet and cold from December to April, dry and hot from May to September and

warm and dry in October and November (Figure 3-3). The climate is characterised by relatively wide annual and diurnal ranges of temperature with restricted rainfall. Temperature reaches 45-50°C in July and 5-10°C in January. The annual precipitation is around 200mm, mostly in winter, and there are seven almost rainless months.

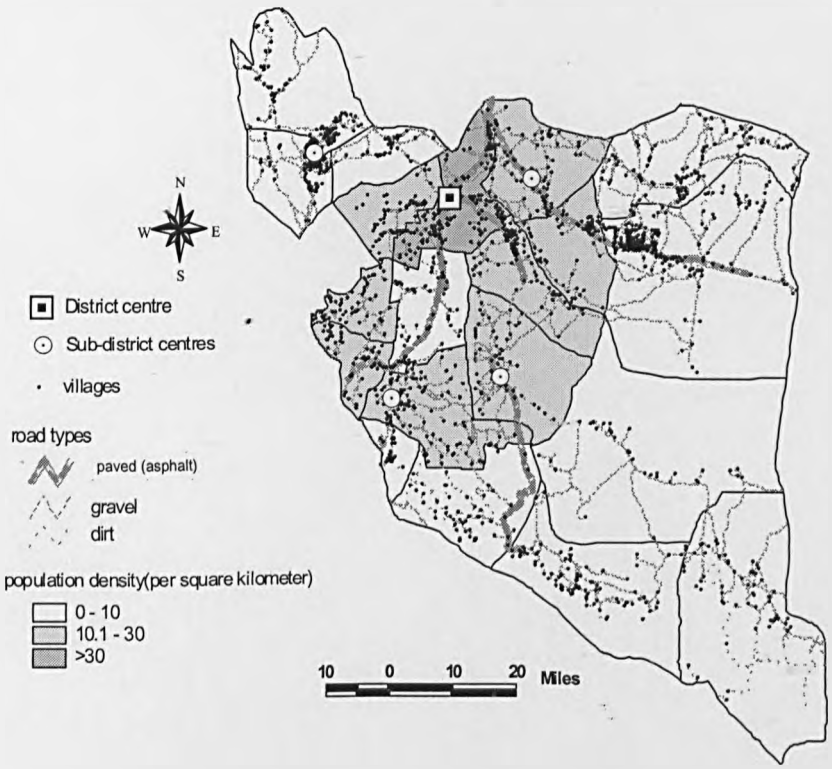

Figure 3-2: Map of Kahnooj district, showing villages, roads and population density by subsubdistrict

The dominant vegetation is scattered bush and scrub, with trees almost always occurring singly. Less than 8% of the district area is used for agriculture purposes and the main crops are date and citrus (Figures 3-4 and 3-5). Deep wells are the main source of water throughout the year. The

water table is around 30-50m but in some parts, it falls sharply to around 100m. Marshes are formed in the wet season but dry out more or less completely at the beginning of dry and hot season.

Altitude varies between of 350m above sea level (in the east) and 2000m (in the south). There is a salt and gravel flat with sand dunes, and a small seasonal lake in east. In the south, there is a mountain chain with rock (Figure 3-6). There are scattered mountains in the northeast and northwest of the district.

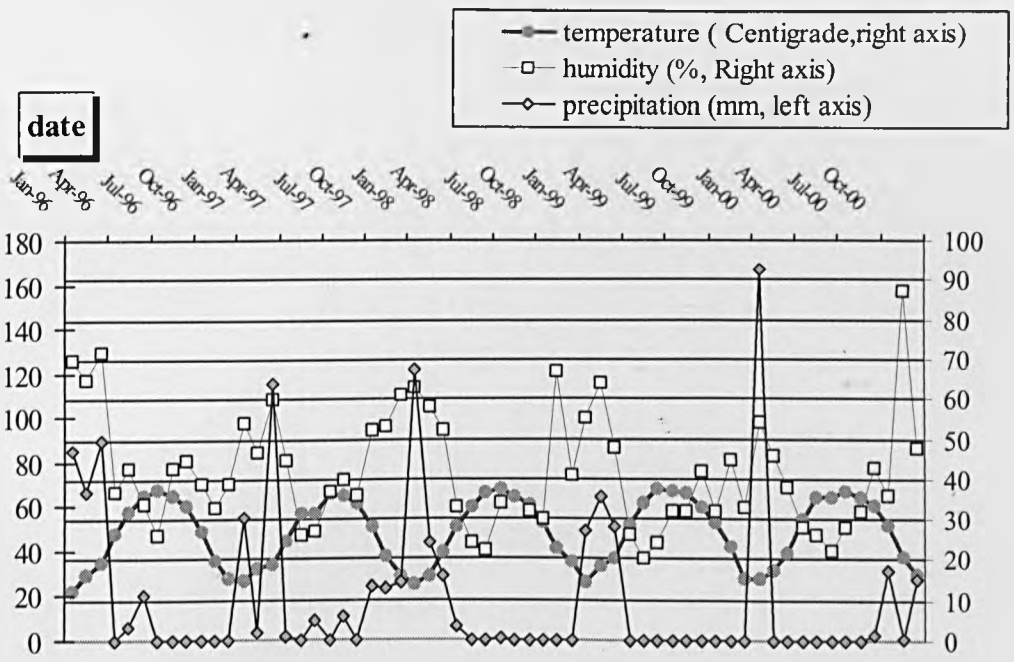

Figure 3-3: Meteorological data of Kahnooj 1996-2000 (Meteorology centre of Kerman Province)

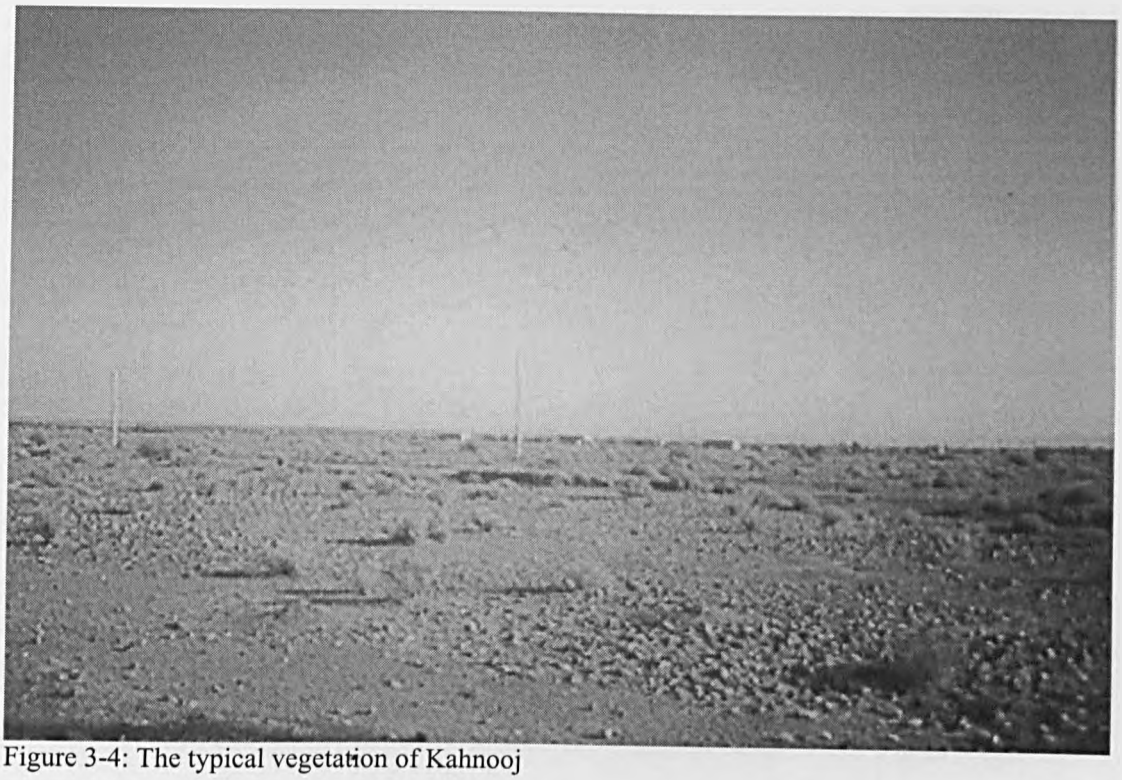

Figure 3-4: The typical vegetation of Kahnooj

Figure 3-5: The main agriculture crops of Kahnooj

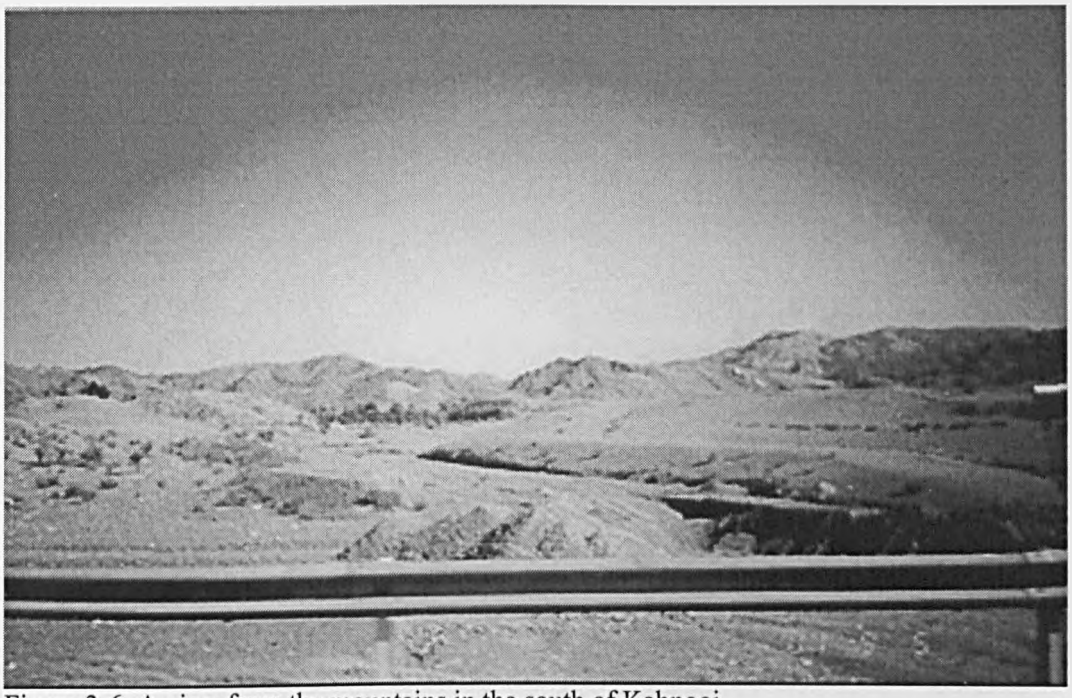

Figure 3-6: A view from the mountains in the south of Kahnooj

3.2.2.2. Population

Kahnooj has a population around 250,000; 86% of which live in rural area. The mean population density is eight people per km^2; the minimum is three in the eastern part of district. Kahnooj has five subdistricts and 20 subsubdistricts (SSDs) (each subdistrict has two to six SSD). In 1994 Kahnooj had just two cities but today has five cities based on national definitions (population more than 7000), and around 600 inhabited villages (Figure 3-2).

Almost all the population belongs to the domestic ethnic group, but there are also estimated to be around 15,000 Afghani refugees in Kahnooj.

More than 60% of people in cities, and 15% in villages, live in brick or cement houses. Most people in rural areas live in clay houses or round huts

made with thatch (Figure 3-7). They live in small and nearly closed groups, practising agriculture and keeping cattle.

Around 19% of the population are nomadic. They move in large groups between this district and adjacent districts following suitable climate and grazing ground. They leave Kahnooj around April to May and come back around August to September, keeping large groups of cattle, and living in huts and temporary tents made of cowhides or grass.

Kahnooj is one of the poorest districts in Iran; 45% of people are illiterate and the annual population growth rate is 1.2% percent. Almost all people are Moslem.

Figure 3-7: Typical accommodation in rural area and mobile people

3.2.3. Description of malaria

3.2.3.1. Malaria in Iran

As the result of extensive malaria control programmes in the last 5 decades, the malaria incidence rate has dropped dramatically. However, malaria is still one of the most common parasitic diseases in Iran (50 to 60 thousand cases per year) and one of the main public health concerns in the south-east of the country *[62]*.

Malaria was endemic in most parts of Iran around 100 years ago. It was estimated that 4-5 million people, out of a population of 13 million, contracted malaria in 1924 *[63]*. Just before the Second World War, Iran started the national malaria control programme, and in 1957 the government included the strategy of malaria eradication. In 1988, the malaria eradication programme became the malaria control programme.

According to the serial reports of Iran to WHO/EMRO (Eastern Mediterranean Region Organisation), there is a decreasing trend in malaria incidence. The 1997 report shows that 10 million (16.4%) of the population lived in initially non-malarious areas, 40 million in areas freed from malaria, 7.5 million in areas of sporadic transmission and just 3.5 million (6%) in areas of constant transmission *[63]*.

Of 3,244,334 blood slides examined in Iran's programme in 1997, 38,766 were diagnosed positive (SPR, Smear Positivity Rate, 1.99%), of which 8,698 (22%) were *P. falciparum* and the remainder *P. vivax* or mixed. 434 cases were hospitalised and 22 deaths were reported.

From a practical point of view, three regions are recognised in the country:

1. The region to the north of the Zagros range, with a population of around 43 million in the west and north-west. API (Annual Parasite Index) was 0.14 per 1000 people in 1997. More than 75% of the cases were imported from other countries (mostly from Azerbaijan) or the endemic areas of Iran.

2. The regions to the south of the Zagros, the central part of Iran, with around 15 million population. API was 0.18 per 1000 population and 48% of the positive cases were imported.

3. The south-east corner, which consists of Sistan and Baluchestan (near to the Pakistan and Afghanistan border), Hormozagan and the southern part of Kerman provinces. The total population was around 3 million and was known as "refractory malaria region" with API 8.74 per 1000 population. The problems facing malaria control programmes were: drug resistance of *P. falciparum*; vector resistance to insecticides; low socio-economic level of people; and the importation of malaria, mostly from Afghanistan and, to a lesser extent, from Pakistan *[63]*.

The long-term research of Edrissian on the drug resistance of malaria showed that most resistant cases were imported, mostly from Afghanistan. In addition, during the last decade, resistance to chloroquine has decreased *[62,64-66]*. Although he reported an in vitro resistance rate of *P. falciparum* of up to 70% in some areas, the low rate of clinical treatment failure means that chloroquine can still be considered as the first line of treatment in Iran *[67]*.

P. falciparum and *P. vivax* malaria patients are treated according to the standard WHO protocol *[68];* i.e., 25 mg/kg chloroquine over 3 days. Primaquine is administered in a single dose of 0.75 mg/kg on day 2 with the third dose of chloroquine as a gametocytocidal drug in *P. falciparum* malaria and as a hypnozoitocidal (anti-relapse) drug in *vivax* malaria. In *P. vivax* patients, administration of primaquine is continued at the same dose, weekly up to and including the 8th week or the same daily dose of primaquine for two weeks.

The five main malaria vectors in the south-east of Iran are *An. stephensi, An. culicifacies, An. fluvuatilis, An. superpictus* and *An. d'thali [62].*

3.2.3.2. Malaria in Kahnooj

Malaria is an endemic disease in Kahnooj, and 1,200 to 3,500 malaria cases are diagnosed every year. Nearly 80% of these are infected with *P. vivax,* and between 1 are 4% are mixed infections. There is virtually no *P. malariae* or *P. ovale.* Severe malaria is rare (less than 2% of all cases) and mortality very rare.

The main and dominant *Anopheles* species are *An. culicifacies* (44%), *An. stephensi* (26%), *An. fluvuatilis* (8%), and *An. superpictus* (4%), the first two of which are the main vectors of malaria as well (Provincial report). The malaria situation in Kahnooj is summarised in Table 3-1. The statistics in this table are calculated from blood slides collected by surveillance system, as described in Section 3.2.4. However, due to very low malaria prevalence in the slides from active surveillance, and the coverage of the pubic health

system, it is reasonable to assume that nearly all cases are identified, and so that the number of positive blood slides might be used as a valid estimator of the number of malaria cases in this district.

Table 3-1: The situation of malaria in Kahnooj from 1997 to 1999 (annual reports of malaria, Health Organisation of Kerman province, 2000)

Year	April 1997 to May 1998	April 1998 to May 1999	April 1999 to May 2000
Population	235,297	249,448	251,315
Number of blood slides	369,918	312,491	235,982
% of slides from active surveillance	96.8	95.8	95
ABER[1]	157.2	125.3	93.9
Number of positive slides			
Active (%)	507(37)	1627(48)	828(43)
Passive (%)	871(63)	1780(52)	1056(57)
Total	1378	3407	1924
SPR[2]			
Active	0.14	0.54	0.48
Passive	7.4	13.5	11.7
Total	0.37	1.1	0.81
percent of *P. falciparum*			
Active	26.9	31	14
Passive	24.7	24.8	8.2
Total	25.4	28.5	10.3
API[3]	5.86	13.66	7.66
AFI[4]	1.49	3.89	0.79

1: ABER (Annual Blood Examination Rate) the number of slides examined for malaria parasites per one hundred population per year
2: SPR (Smear Positivity Rate) the percent of positive slides
3: API (Annual Parasite Index) the number of new malaria cases per one thousand population per year
4: AFI (Annual *falciparum* Index) the number of *falciparum* malaria cases per one thousand population per year

3.2.4. Description of health system

In accordance with the national health programme, there are rural health centres (RHC) in large villages (Figure 3-8). Trained-health workers (HW) are selected from the local people to provide primary health care services compatible with national health policy and local health priorities. In addition, there are mobile health workers based in each RHC who cover the marginal small villages and migratory people. Paramount health centres (PHC) are located in the centres of areas and have general practitioners (GP) and lab equipment, and supervise RHCs.

The total budget of the malaria programme was around US$680,000 in 1999 in Kahnooj. There were 96 RHC in the district. HWs, GPs, microscopists, and the vector control group have the main roles in the malaria control programme.

There are active and passive surveillance systems in this district. HWs, in both rural and urban areas, take blood films from all fever cases that attend health centres (passive surveillance). The method of active surveillance differs according to malaria endemicity. Having asked about the history of fever, the health workers collect blood slides from all people with a positive history since last visit. In addition, in highly infected regions, HWs collect blood slides from around 6% of normal population every 15 days. In infected areas, they collect blood slides from 3 to 4% of population every 15 days and in non-infected areas 1% every month. The definitions of highly infected, infected and non-infected regions are not fixed and the health authorities in the centre of province define their criteria every year.

Trained microscopists stain and read the blood slides. They report the positive results to health workers. Also, they report the results of all slides to the district centre monthly.

All suspected malaria cases are referred to general practitioners (GPs), who diagnosis malaria according to physical exam, history, and thin and thick blood smears. In very remote regions, health workers themselves prescribe anti-malarial drugs according to the WHO protocol for confirmed cases. At least one more blood slide is taken from malaria cases between the 3rd and 5th days after the first dose of anti-malarial drug to check the treatment effect.

The private health sector is active in both rural and urban areas, but it does not have access to anti-malarial drugs, and the diagnosis and treatment of malaria is free of charge in the public sector. Furthermore, private doctors are requested officially to refer all suspected cases to the public health system.

Mosquito surveillance is done in 6 posts monthly, with two methods: total catch (spray sheet collection) and hand catch. The frequencies of species, their sex and the result of anatomical dissections are recorded on standard forms. Nevertheless, due to practical considerations and manpower shortage, especially in hot seasons, these data are not reliable.

Data on malaria cases, collected blood slides and vector surveillance are recorded in 5 standard forms. The cases' forms have data on sex, age, type of accommodation, nationality, dates of collection and slide reading, location, and the response to treatment.

Another section, under the supervision of district health system, is responsible for vector control by using insecticides, larvicides and controlling breeding sites. This group sprays houses and uses larvicides twice per year in areas with transmission (both 'highly infected' and 'infected' areas).

An external quality control scheme is in place under the supervision of the provincial health organisation. It rechecks all positive and 10 percent of negative blood slides. The reference lab checks the slides and recognises three types of discrepancies: 1. Errors in classification of positive and negative slides; 2. Misclassification of species; 3. Errors in reporting of gametocyte/asexual *P. falciparum* forms. All microscopists receive feedback about their errors. Also, the provincial health organisation monitors the malaria situation of districts continuously via the malaria statistics.

Figure 3-8: A rural health centre

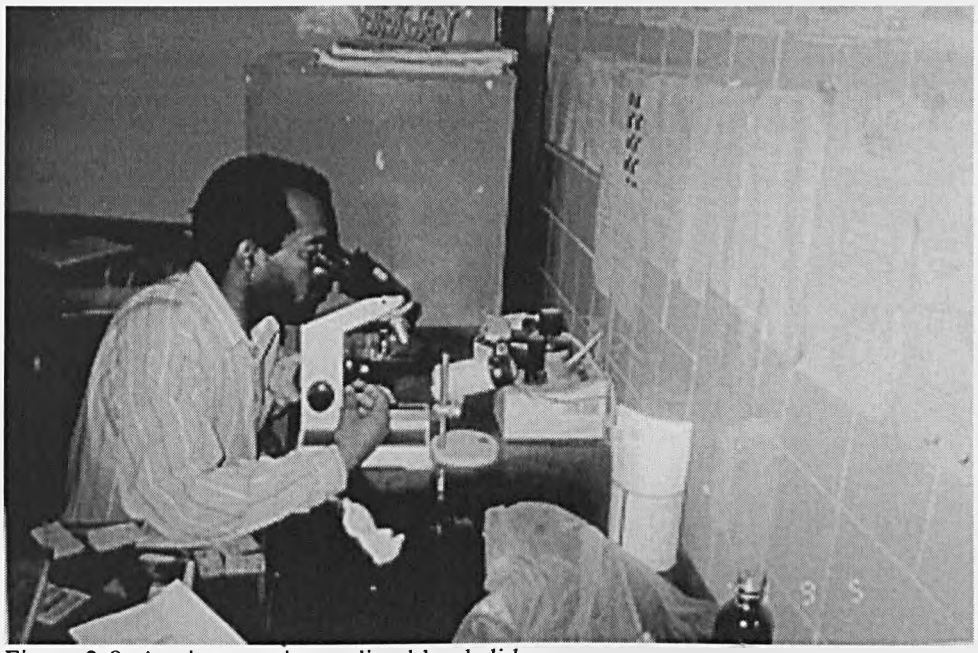

Figure 3-9: A microscopist reading blood slides

3.2.5. Description of meteorology system

The national meteorological network in Iran is linked to the World Meteorological Organisation (WMO) network, which has two types of data collection centres; synoptic and climatological. The synoptic centres have online links to the provincial and national centres and send their data every 3 hours. The synoptic centres are well equipped and measure 18 meteorological variables including wet and dry temperatures, humidity, rainfall, wind speed and direction, and visibility. However, the climatological centres have only basic equipment, and measure wet and dry temperature, humidity and rainfall three times per day at 03.00, 09.00 and 15.00 GMT (06.30, 12.30 and 18.30 in local time), and report their data to provincial centres every week on paper forms.

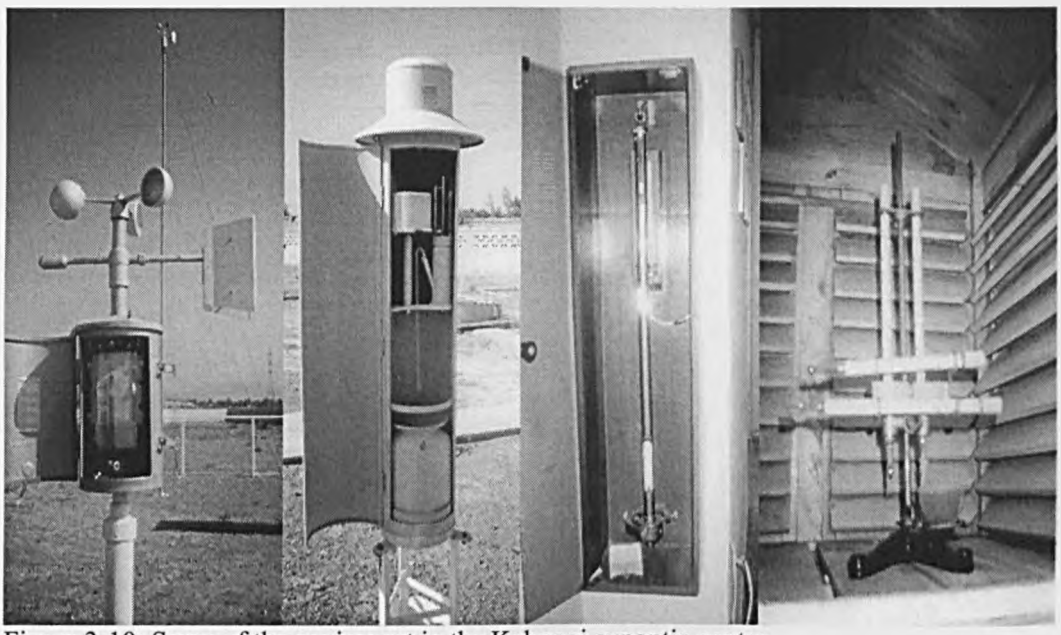

Figure 3-10: Some of the equipment in the Kahnooj synoptic centre

Kahnooj has one synoptic and three climatological centres. The synoptic centre was established in 1979 in Kahnooj city (the centre of the district, latitude 27.58°, longitude 57.42° and elevation 469.7m above sea level). The two climatological centres located in south and centre of Kahnooj district were established before 1980, while the one in southeast of the district was established in 1995. Only data from the synoptic centre will be used here due to missing data from the climatological centres.

3.3. Methods of data collection and analysis

3.3.1. Malaria data

The surveillance forms were the main data source. The frequency of malaria in slides from active surveillance is less than 0.5% (Table 3-1), and almost all cases are enrolled in the surveillance forms (Section 3.2.4); therefore it could be reasonable to calculate risk of disease based only on the surveillance data.

Copies of the original monthly forms for each area were collected from the health organisation of Kahnooj district. These forms had: names of cases and their parents; age; sex; type of surveillance (active, passive); type of accommodation (permanent or temporary); nationality (Iranian or Afghani); location (name of village); and the date of taking and reading blood films. For this study, case data from the 21st of March 1994 (Iranian New year) to the end of 2001 were collected.

Using Epi-Info, the data were double entered. The names of cases and their parents, age, sex and living place (village) were used to trace repeated episodes during the study period.

In order to link the malaria data to remote sensing data, the number of species-specific cases in 10 day (dekad) and one month periods were computed. Remote sensing data contain the mean vegetation index and land surface temperature in each dekad or month (Section 3.5). Therefore, the cumulative number of cases and means of ground climate and vegetation data were computed in each dekad and month.

3.3.2. Demographic data

The populations of villages, subdistricts and SSD were obtained from the statistics of the local health organisation. Their age distributions were estimated according to the national census, which is done every 10 years, most recently in 1996. Village populations for each month were estimated using the population growth rate based on census data in 1997.

3.4. Statistical models

3.4.1. Malaria risks

In this study, dekad, month and annual malaria risks were computed per 100,000 population. By definition, malaria risk was computed as the number of cases divided by the population in each time period.

3.4.2. Mean-median smooth

Smoothing is an exploratory data analysis technique for making the general shape of a series apparent. In this approach, the observed data series is assumed to be the sum of an underlying process (smooth) and of an unsystematic noise (rough) component. Smoothing values (z_t) are obtained by taking mean and/or median of each point in the observed data (y_t) and a few of the points around it. The number of points used is called the span of the smoother.

Since a single smooth usually cannot adequately separate the smooth from the rough, multiple smoothers should be applied in sequence.

In this analysis, the temporal variations of *Plasmodium* spp risks were smoothed using mean-median of span three dekads (Section 3.7.4.2).

3.4.3. Poisson regression method

In all models, the number of malaria cases is the dependent variable. This can only take non-negative integral values and can be modelled by the Poisson distribution. Given a rate λ per unit of time, the probability of observing y cases in a unit time period is then given by Equation 3-1:

$$P(y/\lambda) = \frac{e^{-\lambda}\lambda^{y}}{y!} \qquad \text{Equation 3-1}$$

λ could be predicted using above equation, after taking into account the values of explanatory variables X_1-X_n.

$$\text{Log } (\lambda) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_n X_n \qquad \text{Equation 3-2}$$

where $P(\lambda)$ is the expected number of malaria cases in each time period (X_1 to X_n are the predictors, and β_1 to β_n are the regression coefficients for those predictors.

In this study, seasonal and annual variations, autocorrelation between numbers of malaria cases in consecutive time bands, temperature (ground and remote sensing data), relative humidity, annual rainfall, vegetation index and altitude were used as explanatory variables. The seasonality was modelled using a sinusoidal transformation of time. The annual variability was modelled based on linear and quadratic effect of year. Using quadratic

and multiple regression methods, the optimal model in terms of predictions was constructed.

Using Poisson models, the numbers of cases were modelled based on the explanatory variables, with the log population as an offset (the variable which specified the amount of exposure) (Section 3.7.4).

3.4.4. Fractional polynomial model

Fractional polynomial (FP) regression models provide a flexible parametric method for modelling curved relationships by using few parameters. A FP model extends ordinary polynomial model by including non-positive and fractional powers. A polynomial of degree m may be written as

$$\beta_0 + \beta_1 x + \beta_2 x^2 + \ldots\ldots + \beta_m x^m \qquad \text{Equation 3-3}$$

where as a FP of degree m has m integer and/or fractional powers $p_1 < \ldots < p_m$,

$$\beta_0 + \beta_1 x^{p_1} + \beta_2 x^{p_2} + \ldots\ldots + \beta_m x^{p_m} \qquad \text{Equation 3-4}$$

where for a power p

$$x^p = \begin{cases} x^p \ldots\ldots if \ldots\ldots p \neq 0 \\ \log x \ldots\ldots if \ldots\ldots p = 0 \end{cases}$$

The permitted powers are restricted to the set [-2, -1, -0.5, 0, 0.5, 1, 2, 3]

This family of FP functions may be extended in a mathematically natural way to include repeated powers. An FP of degree m with exactly m repeated powers of p is defined as:

$$\beta_0 + \beta_1 x^p + \beta_2 x^p \log x + \ldots\ldots + \beta_m x^p (\log x)^{m-1}$$

3.4.5. Cox regression method

Survival analysis examines and models the time it takes for events to occur. It typically examines the relationship of the survival distribution to covariates.

Cox regression models the outcome when the values of some explanatory variables change over time. In contrast to log linear models, Cox regression leaves the baseline hazard function unspecified: In other words, an important characteristic of this model is that both the individual and the baseline rates vary with time while their ratio is assumed to remain constant. The general Cox regression equation is:

$$\lambda(t;i)= \lambda(t;0).\theta_i \qquad \text{Equation 3-5}$$

where θ_i represents the hazard ratio which compares $\lambda(t;i)$ (the hazard rate in group i in time t) with $\lambda(t;0)$ (the hazard rate in baseline group in time t).

Cox regression was used to model the rate of secondary attack (Section 3.7.3). Since these rates highly depend on the gap between two attacks, Cox regression was the most appropriate method.

3.4.6. Gaussian approximation

With a Gaussian approximation to the count data, linear regression models could be used. One type of Gaussian approximation is the linear regression model with log transformed dependent variable *[69]:*

$$E(\log(Y+c))=\beta_0+\beta_1X_1+\beta_2X_2+\ldots.+\beta_nX_n. \qquad \text{Equation 3-6}$$

Although some papers showed that Poisson and Gaussian models usually give similar results, particularly with relatively large numbers of cases per unit of time *[70]*, the more sophisticated Poisson model will be used in this research.

3.4.7. Modelling temporal variation

There are three methods to deal with systematic variation of the outcome and predictors with time, in other words auto-correlation *[70]:*

1. Smoothing, this is an appropriate method for Gaussian data, and also count data with a large number of events per each time period. Using moving average methods, such as kernel smoothing, with Poisson regression is straightforward:

$$Log\,(E(Y_t))=\beta_0+\beta_1 X_{t1}+\beta_2 X_{t2}+....+\beta_n X_{tn}+log\,(WMA(Y_t) \qquad \text{Equation 3-7}$$

where $WMA(Y_t)$ denotes the weighted moving average.

An alternative method is to use additive models *[70]*

$$Log(E(Y))=\sum_{1}^{i} S_i(X_i) \qquad \text{Equation 3-8}$$

where S_i is a smoothing function of X_i predictors (including time) and is determined by the data.

2. Semi-parametric approach, in which the predictors are divided up into intervals and a cubic polynomial fitted to each interval. This approach cannot fully check the autocorrelation effect.

3. Parametric approach, which uses a sinusoidal term to fit a wave-like pattern of data. The goodness of fit of models with pure sine transformations of time depends on the origin of time (t=0). Therefore, to maximise the goodness of fit, the time phase (ε) should be estimated, which is the time lag between the t=0 in sine transformation and t`=0 based on calendar ($t=t`+\varepsilon$). The following equations show the method which was used to compute the best time phase (ε):

$$\beta(\sin 2\pi(t`+\varepsilon))=\beta(\sin(2\pi t`)\cos(2\pi\varepsilon) + \cos(2\pi t`)\sin(2\pi\varepsilon)) \quad \text{Equation 3-9}$$

A multiple Poisson regression model with "$\sin 2\pi t$" and "$\cos 2\pi t$" as independent variables gives:

$$\beta_1 \sin 2\pi t` + \beta_2 \cos 2\pi t` \quad \text{Equation 3-10}$$

By comparing the Equations 3-9 and 3-10, it can be written

$$\beta_1 = \beta\cos 2\pi\varepsilon \quad \text{Equation 3-11}$$

$$\beta_2 = \beta\sin 2\pi\varepsilon$$

so that:

$$\beta = \sqrt{(\beta_1^{\,2} + \beta_2^{\,2})} \qquad \text{Equation 3-12}$$

and

$$\varepsilon = \frac{1}{2\pi} \text{Arcsin} \sqrt{\frac{\beta_2^{\,2}}{(\beta_1^{\,2} + \beta_2^{\,2})}} \qquad \text{Equation 3-13}$$

The number of malaria cases in each stratum was low if the data were subdivided by spatial and temporal factors. Furthermore, the results of a pilot study indicated that a parametric approach with sinusoidal transformation was a valid approach. Therefore, method three; parametric method, was used to model the temporal variations.

Simple graphs showed a decreasing time trend in the risks of species-specific malaria. Therefore, between years variation was modelled using linear and quadratic effects. In order to deal with the co-linearity between the linear and quadratic terms, 1993 was subtracted from the original year variable; i.e., the first year of study (1994) was recoded to year one.

Malaria risk in consecutive time bands has autocorrelation (Section 3.7.4.5). The above statistical techniques were applied to the data using time series commands in Stata. More theoretical details can be found in papers written

by Dominici et al. (2000) *[71]*, Schwartz et al. (1996) *[70]* and Katsouyanni et al. (1996) *[69]*.

3.4.8. Accuracy of models

The data were grouped by month or ten day period; then the grouped malaria data were randomly allocated to modelling (75%) and checking (25%). To do so, the records were sorted by date, then the first record of the checking part was selected based on a random number between one and four (which was two); and every four the record was included in the checking part (records number two, six, ten and so on). The parameters were estimated from the modelling data. Model accuracy was checked by comparing the fitted and observed values in the checking data adjusted for spatial and temporal clustering.

The main objective of this study was to assess the feasibility of an early warning system. The model accuracy was assessed in checking part because predictions from a given set of parameters are better when they are used in the data from which were derived.

Using the forward selection method, the variables were entered into Poisson models. The simplest model estimated the number of cases in each time band based on seasonality (sine and cosine terms). Then, step by step, other explanatory variables were added to the models and their influences assessed according to their impact on the pseudo R^2, and their statistical significance using likelihood ratio tests.

The most common assessment of overall model fit in logistic regression is likelihood ratio test, which is simply the chi-square difference between the null model (i.e., with the constant only) and the model containing one or more predictors or between two nested models.

Nonetheless, in logistic regression, there is no true R^2 value as there is in least square regression. However, because deviance is analogous to residual value in the regression analysis, one can approximate an R^2 based on lack of fit indicated by the deviance (-2LL).

$$R^2_{\text{logistic}} = 1 - \frac{-2LL_{reduced} - 2LL_{full}}{-2LL_{reduced}} = 1 - \frac{G}{-2LL_{reduced}} \qquad \text{Equation 3-14}$$

where the reduced model is the logistic model with just the constant and the full model contains all the predictor in the model.

Having selected the best combination of explanatory variables in the modelling part of data, model accuracy was assessed in the checking part. Residuals are defined as the differences between observed (y_t) and predicted values (z_t).

The Sum of the Absolute Residuals (*SAR*) was defined as:

$$\sum_{t=1}^{n} | y_t - z_t | \qquad \text{Equation 3-15}$$

where n is the total number of time bands

This is the most straightforward statistic for measuring model accuracy. It measures how far the total estimated values are from the observed values. From a practical point of view, it helps health policy makers to know how

many cases would be under or over estimated by the model. Nevertheless, to differentiate the amount of over and under-estimations, two different indices were defined: Sum of Over-estimations (SO) and Sum of Under-estimation (SU) as follow

$$SO = \sum_{t=1}^{n} z_t - y_t \text{ if } y_t < z_t$$
$$SU = \sum_{t=1}^{n} y_t - z_t \text{ if } y_t > z_t$$

Equation 3-16

SAR, *SO* and *SU* depend on the duration of follow up. To deal with this problem, the Average Annual number of Over and Under-estimation (*ASO* and *ASU*) was computed based on the following formulae:

$$ASO = \frac{SO.k}{n}$$
$$ASU = \frac{SU.k}{n}$$

Equation 3-17

where k is the number of time bands per year.

The square root of the sum of squared residuals is another index which can be used to assess model accuracy. It is defined as:

$$\sqrt{\left(\sum_{t=1}^{n}(y_t - z_t)^2\right)}$$

Equation 3-18

Form the statistical point of view, this statistic has some advantages compared to the other explained indices such as SAR, particularly in the least square regression models; however, it is more difficult to interpret.

Finally to show the Relative Over and Under-estimations, *RSO* and *RSU*, the *SO* and *SU* were divided by the total number of observed cases:

$$RSO = \frac{SO}{\sum_{t=1}^{n} y_t}$$

$$RSU = \frac{SU}{\sum_{t=1}^{n} y_t}$$

Equation 3-19

Because this project aims to test the feasibility of an early warning system in Kahnooj, models and their results should be as clear as possible to health policy makers and it would be more practical to show the accuracy of models with relatively simple measures. Hence, the accuracy of the models was estimated using the four statistics *ASU, ASO, RSU* and *RSO.*

3.5. GIS and RS data

3.5.1. Electronic maps and field data

Various digital datasets, including existing data available from the Health Ministry of Iran, were used to model the spatial distribution of malaria cases and their relation to meteorological variables, altitude and land use/cover in three spatially distinct levels: village (high resolution), SSD (middle resolution) and district (low resolution).

Electronic maps of Kahnooj contain the borders, roads, villages and cities. The map scale was 1:50,000 in Arcview format. The maps were developed

by the national geographical organisation in Iran and were updated in 2000 and generated based on aerial photographs.

Using a GPS, the accuracy of co-ordinates of villages and roads in GIS files were assessed. Latitude, longitude and altitude of 20 points (mostly branching point of roads) and 14 villages were recorded; these points and villages were scattered geographically in the district. The maximum difference between the recorded co-ordinates of points and GIS files was less than 30m, while the maximum difference for villages was 300m. The greater errors in villages were mostly due to the locations within villages which were surveyed by GPS; the buildings in villages were scattered mostly in an area of 2-3 km^2, depending on the location of the selected point for GPS, the result might have around one kilometre difference with the points in the electronic maps. Based on the above findings, it seems that the electronic maps are accurate enough to be used in this study.

In addition, the ground data of land cover/use were matched with RS imagery. However, due to tiny differences in vegetation indices and little association between malaria risks and vegetation index, the results of these classifications are not reported.

3.5.2. Software

In this analysis two GIS programs were mainly used: Arcview 3.2 (Environmental Systems Research Institute, Inc.) and Idrisi 32 (the George Perkins Marsh Institute, Clark University). In addition, the Landsat data was processed using ENVI software (version 4, RSI, Boulder, Colorado). Arcview is a vector digitising GIS software, which captures each point as a

pair of (x, y) co-ordinates with extensive spatial analysis features. Idrisi is based on a raster data model, which converts the maps into grid composed of individual pixels. All satellite images are in raster format.

Having processed images in Idrisi, the extracted data were transferred to Stata for modelling and Arcview for visualisation. Idrisi has powerful facilities to manipulate and extract image data.

3.5.3. Remote sensing images

Land use/land cover was derived from supervised classification of digital Landsat ETM+ imagery. All other data were obtained from secondary sources. Modelling using land cover/use was initially restricted to southern Kahnooj. Landsat data has 30m spatial resolution (path 159, row 041, 7th January of 2001, path 158, row 041, 31st of December 2000 and path 158, row 042, 31st of December 2000). Also, another data set with 8km spatial and 10 days temporal resolution was obtained from NOAA-AVHRR from 1990 to 2001. This data set has NDVI (Normalised Differentiate Vegetation Index) and LST (Land Surface Temperature). The first and second data sets were used to model the frequency of malaria by land cover/use, and NDVI and LST respectively.

3.5.4. Processing

Basic radiometric correction of the Landsat data was carried out using ENVI. Raw data were converted to reflectance for each of the seven spectral bands on the basis of the red and near infrared channels (channels 3 and 4 in the case of TM), using the following standard formula:

$$NDVI = \frac{(NIR - RED)}{(NIR + RED)}$$

NDVI images were subsequently exported to Idrisi for further spatial analysis.

Dekadal NDVI data at an 8x8km resolution were provided by Dr. S.I. Hay for the period of January 1981-September 2001. These data were derived from NOAA-AVHRR satellite data using standard procedures. *[48]* Village specific time series for NDVI were calculated in this instance by overlaying village coordinates (on each NDVI decadal image in turn) using ENVI software.

The altitude was extracted from DEM (Digital Elevation Models) in one square kilometre resolution. The images were downloaded from the website of National Imagery and Mapping Agency of United State of America (http://geoengine.nima.mil/). Using ENVI version 4 (Environment for Visualising Images), the images were converted to Idrisi format. Then, in Idrisi, the average of elevation around villages were computed (buffer zone), using sensitivity analysis the best buffer zone was defined.

3.6. Ground climate data

The climate data were collected in the synoptic centre in Kahnooj City. Mean daily temperature and relative humidity were computed as the average of minimum and maximum values.

3.7. Results

To assess the accuracy of microscopy on detection of *Plasmodium* spp in Kahnooj, Section 3.7.1 compares the results of microscopy and Polymerase Chain Reaction (PCR) method in a small sample. Section 3.7.2 describes the overall situation of malaria in Kahnooj, and assesses the relationships between the frequencies of species and demographic variables. Using time series analysis, Sections 3.7.4 and 3.7.5 explain the temporal and special distributions of cases. Sections 3.7.5.2, 3.7.5.3, and 3.7.5.4 assess the goodness of fit of models based on remote sensing variables; in these models the numbers of cases in each 10 days in each village are linked to the vegetation indices, land surface temperature and altitude. The accuracies of models based on ground and remote sensing climate data are checked in Section 3.7.6. In the last Section (3.7.7), the local transmission risk was modelled based on climate and remote sensing data.

3.7.1. The accuracy of microscopy results

3.7.1.1. Introduction

Light microscopy has historically been the mainstay of the diagnosis of malaria in Iran (Section 3.2.4). The national health policy of Iran dictates that the clinical diagnosis of malaria disease should depend on visualisation of parasites by light microscopy of Giemsa-stained blood smear in febrile cases. This procedure is cheap and simple, but it requires well-trained personnel.

Microscopists are expert in Iran, and there is a quality control programme which supervises the microscopy results (Section 3.2.4). However, there is only one published paper on the accuracy of microscopy results compared to PCR results in Iran, which was written by Zakeri at al. (2002) *[72]*. They double checked the microscopy results of 120 fever patients in the Chahbahar district of Sistan and Baluchestan Province in south-eastern Iran. They found microscopy had more than 95% specificity in the detection of *P. falciparum* and *P. vivax*. However, this technique missed around 75% of mixed-species malarial infections.

While the results of the Chahbahar study are useful, several factors may obscure the true accuracy of microscopy for the detection of *Plasmodium* spp. This is especially true when comparing the results found in Chahbahar with those expected for Kahnooj. For instance, the endemicity of malaria, particularly *P. falciparum*, is higher in Chahbahar than in Kahnooj. In addition, Chahbahar is located in a very remote area close to the Pakistan border and has an even poorer population than Kahnooj. This is an important factor considering that the entire Sistan and Baluchestan province has the most under-developed health system in Iran. On the other hand, Kahnooj has an efficient malaria surveillance system, with trained personnel. Therefore, it could be expected that the microscopy results in Kahnooj are more accuracte than in Chahbahar. However, to date there has not been any objective evidence to show the accuracy of their microscopy findings.

PCR is a useful tool to validate the effectiveness of light microscopy in the detection of malarial parasites. PCR has greater sensitivity and specificity than light microscopy *[73-76]*, particularly in situations of low-level

parasitaemia *[77]*. Furthermore, it is a more powerful technique to detect mixed infections of malarial species *[78]*.

3.7.1.2. Materials and methods

During fieldwork in summer 2002, three highly endemic villages were selected in the north-west, centre and south-east of Kahnooj. In August and September 2002, a systematic sample of fever patients who sought treatment at the health centres of these villages was included. The sampling was based on the days of week: only patients who came to health centres in the 2nd and 4th of Iranian working days (Sundays and Tuesdays) were assessed.

First, informed consent from patients or their guardians was obtained, as was information on the age, sex, duration and symptoms of disease. Next, finger-prick blood samples were collected: thick blood slides were prepared for microscopical observation and, for comparison; three blood dots were dropped directly on filter mats for the PCR assay.

For the thick blood slide analysis, all the blood slides were air-dried, fixed in methanol and then stained in Giemsa for 15–30 minutes; a 1:10 dilution of Giemsa (pH 7.2) was used. The stain was washed off with tap water. Next, the slides were read by microscopists with the routine methods, i.e., oil immersion lens at x1,000 magnification for at least 100 oil immersion fields; also, an expert microscopist in the reference laboratory in the centre of the province re-checked the slides blindly.

For the PCR assay, extraction of parasite Deoxyribonucleic Acid (DNA) was carried out using a chelex extraction method described by Walsh et al (1991)

[79]. Briefly, the blood samples on filter mats were thawed, and the parasite DNA was extracted by boiling with 20% chelex resin after the samples were left overnight in 1xPBS/0.5% saponin. DNA samples were processed by PCR to amplify species-specific sequences of 18s subunit ribosomal ribonucleic acid (18ssrRNA) genes of *P. vivax* and *P. falciparum*. All positive samples based on microscopy and ninety of the slide negative samples were also extracted on a 96-well plate. Fifteen µl DNA was pooled together once into a 'row' and once into a 'column' from each negative sample. Each pool was cleaned by phenol/chloroform and ethanol, precipitated and re-suspended in 15µl.

Using Epi-info 6, the data were analysed and the Sensitivity, Specificity, Positive and Negative Predictive Values (PPV and NPV) and their 95% confidence intervals were computed. These indices are for slide reading, taking PCR as the gold standard.

3.7.1.3. Findings

A total of 124 patients were included in this study; the mean and standard deviation of age were 20.2 and 16.7 years respectively (with minimum and maximum of 6 months and 77 years); 60 were male (48.4%) and 64 female (51.6%). All patients had a history of fever in last 48 hours prior to seeking treatment; 43% had pain, of which headache was the most common form; 38% had shivering, and 6% vomiting. (Table 3-2)

Table 3-2: Distribution of sex, age and location; and history of symptoms among 124 subjects

	Frequency	percentage
Sex		
Male	60	48.4
female	64	51.6
Location		
North-West	33	26.6
Centre	67	54.0
South-East	24	19.4
Age (year)		
<10	38	30.6
10-20	43	34.7
21-40	21	16.9
>40	22	17.7
History symptoms		
Fever	124	100
Shiver	47	38
Pain	53	43
Vomiting	7	6

Based on microscopist reports, ten patients were infected with *P. vivax* and none with *P. falciparum* (Table 3-3). These results were exactly the same as the referral laboratory reports.

The PCR results are shown in Figure 3-11. This figure illustrates that all of the positive slides for *P. vivax* based on microscopy were also detected as positive by PCR; none of these patients had mixed infections. On the other hand, three of the negative slides for *P. vivax* based on microscopy were detected as positive by PCR (samples in row A and columns 2, 4 and 9).

The sensitivity and specificity of microscopy in the detection of *Plasmodium* spp infection were 77% (95% CI: 46%-94%) and 100% (95% CI: 95%-100%), correspondingly. Also, the estimated positive and negative predictive values were 100% (95% CI: 66%-100%) and 97% (95% CI: 91%-99%), respectively (Table 3-3)

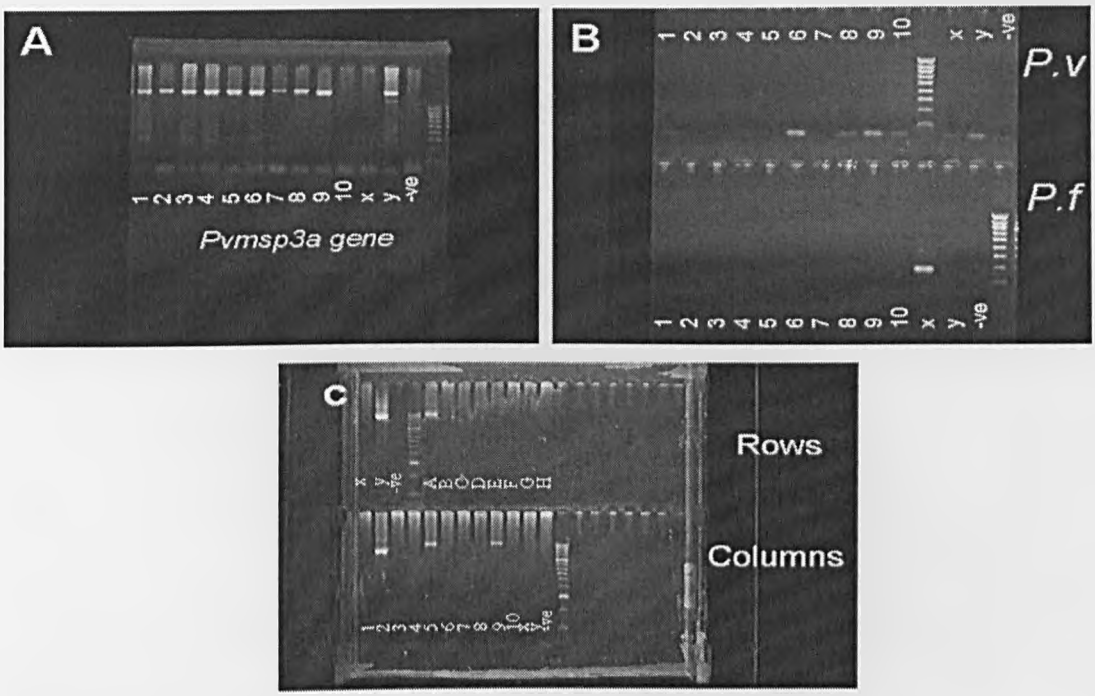

Figure 3-11: Gel nested PCR products; clinical specimens using species-specific oligonucleotide pairs for A (Pvmsp3a gene), and B (species-specific amplifications) in *P. vivax* slide positive samples, and C (Pvmsp3a gene) in pooled slide negative samples, which shows positive PCR in samples in row A and columns 2, 5 and 9. Markers are 100 bp ladders. x: 3D7, y: v97007 (a *P. vivax* gene extracted from Venezuela samples); and -ve: negative control

Table 3-3: The accuracy of microscopy in detection of *P. vivax* among fever patients in Kahnooj

PCR / Microscopy	positive	negative	total
Positive	10	0	10
negative	3	97	100
total	13	97	110

Sensitivity= (10/13).100=77%
Specificity= (97/97).100=100%

PPV[1]=(10/10)*100=100%
NPV[2]=(97/100)*100=97%
1: Positive Predictive Value
2: Negative Predictive Value

3.7.1.4. Discussion

In contrast to the sensitivity and specificity, the predictive values are dependent on the disease prevalence. The blood samples were taken in the

peak of malaria transmission season in highly endemic villages. Therefore, it could be expected that the NPV was higher in the whole district; i.e., negative blood slide based on light microscopy in the current setting can rule out *Plasmodium* spp infections in febrile cases with at least 97% precision, and positive slides confirm infection with 100% precision. Hence, from a medical point of view, the result of microscopy is a very accurate tool in diagnosis of malaria.

The three false negative samples were taken from two females aged 45 and 55 years, respectively, and one male aged 18 years; two of them had a history of fever with vomiting, and the third one had only a history of fever. The blood slides of these three patients were re-read by an expert in the reference laboratory, and 300 fields were assessed per slide. No parasites were detected in any of these slides. Therefore, the discrepant results may either be due to very low levels of parasitaemia and false negative microscopy, or false positive results from the PCR as a result of cross-contamination.

Assuming that the discrepancy between microscopy and PCR is the result of the lower sensitivity of the former method, reassessment of the slides was negative. Therefore, it could be concluded that these slides were from patients with very low levels of parasitaemia. If this is the case, it should be noted that such low-level parasitaemia can rarely be diagnosed by ordinary light microscopic methods.

Boisier et al. (2002) showed an association between the level of parasitaemia and fever in a seasonal transmitted area; i.e., the signs and symptoms of malaria are not common in very low levels of parasitaemia *[80]*,

which supports others findings even in highly endemic places *[81-83]*. Therefore, the presence of fever in these three subjects may be due to other febrile diseases which are common during the summer in Kahnooj, e.g. gastroenteritis.

Malaria disease is defined as fever plus positive blood slide with or without other signs or symptoms in endemic areas of Iran. In this project, malaria disease, not infection, is the outcome. Therefore, for the purpose of this study, it can be implied that these three false negative results do not change the accuracy of microscopy results. In other words, since even after very extensive examinations the slides of these three subjects remained negative by microscopy, these patients do not meet the eligible criteria to be classified as malaria patients; therefore, the estimated sensitivity, specificity and predictive values show the accuracy of microscopy in diagnosis of malaria infection, not malaria disease.

Although the sample size might be a point of concern in extrapolating the result of this study, it can be concluded that the light microscopy had satisfactory accuracy in detection of malaria infection and particularly disease in Kahnooj.

3.7.2. Overview of malaria data

Between March 1994 and March 2002, 18,268 malaria attacks were recorded in Kahnooj, of which 12,337 (67.5%) were infected with *P. vivax*, 5,858 (32.1%) with *P. falciparum*, and 73 (0.4%) had mixed infections.

Sixty percent of attacks (10,680) were detected by passive surveillance, and 40% (7,150) were identified by active surveillance.

Table 3-4 relates the frequencies of different species with sex, type of accommodation, nationality, type of surveillance, and age. Compared to *P. vivax*, *P. falciparum* was relatively less common in males, those who lived in temporary accommodation, Afghani people, and also cases detected by passive surveillance. All these differences are statistically significant, although the magnitudes of differences are not considerable.

Table 3-4: Description of malaria cases in Kahnooj between March 1994 and March 2002

Number (percent)	*P. vivax*	*P. falciparum*	Mixed	Total	χ^2(df) p-value
Sex					
Male	6,788(55.1)	3,095(52.8)	49(67.1)	9,932(54.4)	12.8 (2)
Female	5,539(44.9)	2,763(47.2)	24(32.9)	8,326(45.6)	0.002
Accommodation					
Permanent	11,781(97)	5,620(99.1)	70(98.6)	17,471(99)	70.9 (2)
Temporary	350(2.9)	50(0.9)	1(1.4)	401(2.2)	<0.001
Surveillance					
Active	4,563(37.7)	2,541(44.9)	46(63.9)	7,150(40.1)	101.2 (2)
Passive	7,541(62.3)	3,113(55.1)	26(36.1)	10,680(60)	<0.001
Nationality					
Iran	11,758(95)	5,800(99)	70(98.6)	17,628(97)	161.4 (2)
Afghanistan	569(4.6)	56(1)	1(1.4)	626(3.4)	<0.001
Age					
<5	2,008(16.3)	957(16.3)	7(9.6)	2,972(16.3)	
5-14	5,017(40.7)	2,377(40.6)	42(57.5)	7,436(40.7)	
15-29	3,326(27)	1,656(28.3)	19(26)	5001(27.4)	P<0.001[1]
=30	1,972(16)	865(14.8)	5(6.8)	2,842(15.6)	
Mean (se)	16.66(0.13)[2]	16.07(0.18)[2]	13.3(1.19)	16.46(0.11)	

1: p-value of ANOVA test on natural logarithm of age
2: The result of Tukey HSD test showed a significant difference only between the mean age of *P. falciparum* and *P. vivax* infected cases

Having used the Kahnooj census data in 1997, the effects of sex, age, nationality and type of accommodation were assessed as risk factors for malaria disease. Table 3-5 shows that the risk of disease in females was around 14% lower than males. Also, the disease risk was more or less constant in those under 30 years old, while in older people (30 or more year old) the risk of disease was halved.

Surprisingly, inhabitants in temporary accommodation and particularly Afghani people had less chance of disease (accommodation RR=0.91, nationality RR=0.52). Since many Afghanis lived in temporary houses, the accommodation RR might be confounded by nationality.

Nevertheless, lower risk of reported malaria disease among Afghanis might be explained by either information bias or under-estimation of Afghani cases. Most Afghanis live in Kahnooj illegally, and often do not disclose their nationality; therefore, it could be expected that their nationality was not always recorded accurately in health system files. Also, Afghanis usually live in very remote areas, with relatively poor accessibility to health facilities; furthermore, they might have less motivation to get health advice.

In addition, most Afghanis immigrated to seek a job; they might be older than the average Kahnooj population and, due to previous frequent exposures to *Plasmodium* spp, might be less susceptible to the malaria disease.

Table 3-5: The risk of malaria disease, classified by sex, age, nationality and accommodation type

	number of Malaria cases	Population[1]	Disease risk between 1994 and 2002 (per 100 population)	Risk ratio (95% CI)
Sex				
Male	9,932	98,330	10.2	1
Female	8,326	97,950	8.7	0.86 (0.83-0.88)
Accommodation				
Permanent	17,471	191,400	9.1	1
Temporary	401	4880	8.2	0.91 (0.83-1.0)
Nationality				
Iran	17,628	179,936	9.8	1
Afghanistan	626	12,950	4.8	0.52(0.48-0.56)
Age				
<5	2,972	28,571	10.4	1
5-14	7,436	66,316	11.2	1.07(1.03-1.11)
15-29	5,001	48,498	10.3	0.99(0.95-1.04)
>=30	2,842	50,962	5.6	0.56(0.53-0.59)

1: The population are extracted from census data of Kahnooj in 1997

3.7.3. Repeated malaria attacks

This section explores the frequency of multiple attacks within cases, and estimates the risk of the therapeutic failure as the frequency of secondary attack within 30 days after the first attack with the same species (Section 3.7.3.1), and relapse risk of *P. vivax* as the difference between the risk of secondary *P. vivax* attack in those who had primary *P. falciparum* and *P. vivax* infections (Section 3.7.3.2). Although this section does not directly address to the main objective of this chapter which was the feasibility of an early warning system, its results are very important to understand the epidemiology of malaria in Kahnooj.

Multiple disease episodes might be attributable to any of these reasons:

1. Treatment failure: cases might either take insufficient main treatment (chloroquine) to eliminate the blood forms of *Plasmodium* spp (schizonts, merozoites and trophozoites), or their infections might be resistant to the administrated drugs. In this group, the gap between consecutive episodes would be short, a few weeks at maximum, because the parasite was not effectively eliminated from their blood *[84-88]*.

2. Re-infection: *Plasmodium* spp do not generate full protective immunity, particularly in areas with low endemicity; *[89-92]* therefore, cases may be re-infected. Furthermore, due to heterogeneity in exposure risk *[27,93,94]*, it could be expected that cases are generally more exposed to vectors than non-cases, and so could have a greater chance of re-infection.

3. Relapse: this could be a reason only for those who were infected with *P. vivax* in both episodes. It is practically impossible to differentiate re-infection from relapse; therefore, most papers have estimated the joint risk of re-infection and relapse; i.e., they reported the risk of secondary *P. vivax* attack due to either of these reasons *[95-103]*. Nevertheless, some papers reported the risk of secondary *P. vivax* attacks during non-transmission season as an estimation of the relapse risk *[104,105]*.

The joint risk of *P. vivax* relapse and re-infection without any hypnozoitocidal (anti-relapse) therapy is around 25-55% *[98-100,105,106]*. There is controversy about the efficacy of primaquine as anti-relapse

treatment; Rowland and Durrani (1999) showed in his clinical trial in an Afghani refugee settlement in Pakistan that 5-day primaquine therapy did not have any significant effect against of *P. vivax* relapse (52% versus 51% first relapse or re-infection risk in control and treatment arms respectively), but 14-day therapy reduced relapses significantly (49% versus 34% first relapse or re-infection risk, in control and treatment arms respectively). This finding was also supported by a large study in India *[107]*, another clinical trial in Pakistan *[106]* and a small study in Sri Lanka *[97]*. In contrast, Srivastava et al. (1996) *[103]* showed that even a 5-day primaquine regime could decrease the relapse rate; Roy et al. (1977) and Sharma et al. (1990) showed that 5-day therapy could prevent more than 60% of *P. vivax* relapses *[101,102]*.

P. vivax exhibits three main patterns of relapse activity: 1) tropical pattern with short latency (1-3 months) between attacks, 2) temperate pattern with a latent period of 6-14 months, followed by renewed parasite activity in the form of one or more relapses with short intervals between each relapses, 3). intermediate form *[104,105]*. However, there is one case report of a *P. vivax* relapse occurring many years after the primary attack *[108]*.

Among the available data in Kahnooj, there was not a unique variable to link the data of repeated disease episodes within cases. In this study, the records were linked based on case's name, the first and last letters of parents' names, living place and age. Twenty five records (out of 18,268) were excluded from this analysis due to missing fields. Then, the gaps

between two consecutive episodes were assessed, and classified by the detected species in the former and latter episodes.

The 18,268 malaria attacks were recorded in 16,297 persons. 14,799 (91%) of cases had just one, and 1,169 (7.2%) had 2 attacks; less than 2% of cases had more than 2 attacks. The frequency of more than two attacks was low, but, nevertheless, the following results have been adjusted for within person clustering effect. Also, only 73 cases had mixed infections with *P. falciparum* and *P. vivax*, of whom only 23 showed repeated attacks. These 23 mixed infected cases have been excluded from those analyses related to the species in the former and latter attacks.

Table 3-6 illustrates that the gap between two consecutive attacks depends on the species (Pearson χ^2=652.6, df=18, p<0.0001; Kruskal-Wallis χ^2=451.26, df=3, p<0.0001).

Table 3-6: The gap between two consecutive attacks according to the species in the former and latter episodes

Former infection		*P. falciparum* (%)		*P. vivax* (%)		Total
Latter infection		*P. falciparum*	*P. vivax*	*P. falciparum*	*P. vivax*	
The gap between two episodes (months)	=1	169(38.1)	15(4.3)	23(9.7)	28(3.1)	235(12.2)
	1-2	133(29.9)	25(7.1)	44(18.5)	59(6.5)	261(13.4)
	3-5	33(7.4)	22(6.2)	37(15.5)	72(7.9)	164(8.4)
	6-8	11(2.5)	22(6.2)	5(2.1)	79(8.7)	117(6)
	9-11	17(3.8)	40(11.3)	17(7.1)	149(16.4)	223(11.5)
	12-17	26(5.9)	36(10.2)	44(18.5)	134(14.8)	240(12.3)
	=18	55(12.4)	193(54.7)	68(28.6)	387(41.6)	703(36.2)
	Total	444(100)	353(100)	238(100)	908(100)	1943(100)

Pearson $\chi 2=652.6$, df=18, $p<0.0001$; Kruskal-Wallis $\chi 2=451.26$, df=3, $p<0.0001$
Mixed infections and subjects with more than 2 attacks were counted more than once accordingly

In contrast to the other groups, *P. falciparum-P. falciparum* attacks occurred mostly within a short gap, which might provide some idea about treatment failure rate. On the other hand, the gap between two attacks was longer in *P. vivax-P. vivax* group, which gives an impression about relapse rate. The following sections present the results of Cox models to address these issues more precisely.

3.7.3.1. Therapeutic failure

Malaria treatment failure is a general term and is used when the case does not fully respond to anti-malaria drugs. In most such cases, drugs may decrease the blood density of parasites, even to levels undetectable by light microscopic exam, but after a short time the parasite density rises again. WHO defines early and late treatment failure based on the gap between the

date of the first dose of the drug and the date of second parasitaemia. WHO recommends that cases be followed up to 28 days *[109,110]*. Therefore, studies usually report the first month treatment failure risk or rate *[111-114]*. In this analysis, any secondary attack within 30 days was considered as treatment failure. Nonetheless, it should be mentioned that the optimum cut-off would depend on the intensity of transmission; the higher the transmission the more likely an earlier re-infection; and this would in principle affect the optimum cut-off time.

Around 38% of *P. falciparum-P. falciparum* attacks occurred within 30 days, which was much higher than in *P. vivax-P. falciparum* (9.7%) and *P. vivax-P. vivax* (3.1%) (Table 3-6).

To check the treatment failure rates of *P. vivax* and *P. falciparum*, Cox regression models were used. Having stratified for the gap between two consecutive attacks, the number of cases was defined as the number of secondary attacks with the same species as the primary attack; the duration of follow up was measured in person-months; the rates were adjusted for within person clustering effect and season.

Among 566 person-months follow-up, 145 *P. falciparum* attacks were recorded with less than one month gap; i.e., the *P. falciparum* treatment failure rate was 256.2 per 1000 person-months (95% CI: 217.7-301.5). The corresponding rate for 1-3 months gap was 146.6 (95% CI: 123.2-174.3); after that, for longer gaps, the rates dropped sharply to less than 30 (Figure 3-12). Among 851 person-months follow-up, 26 *P. vivax* attacks were recorded with less than one month gap, the estimated *P. vivax* treatment failure was 30.5 per 1000 person-months (95% CI: 20.8-44.8). The variation

of *P. vivax* rates was not compatible with those of *P. falciparum*, which has been explained in details in Section 3.7.4.3.

The rate ratios of sex, age group and nationality for *P. falciparum* treatment failure were not significant; however, the corresponding rate ratio for type of surveillance was 2.0 (95% CI: 1.4-5.9). In other words, the rate of recorded *P. falciparum* treatment failure in active surveillance was around twice that in passive surveillance. This difference could be explained simply by the follow-up method of cases in Kahnooj. The health workers should take at least two follow up blood slides from *P. falciparum* cases actively (Section 3.2.4). Therefore, a *P. falciparum* treatment failure was more likely to be detected in active than passive surveillance.

The rate of *P. vivax* treatment failure was not significantly associated with sex, age-group, nationality or type of surveillance, i.e., none of these factors could be counted as risk factors for *P. vivax* treatment failure.

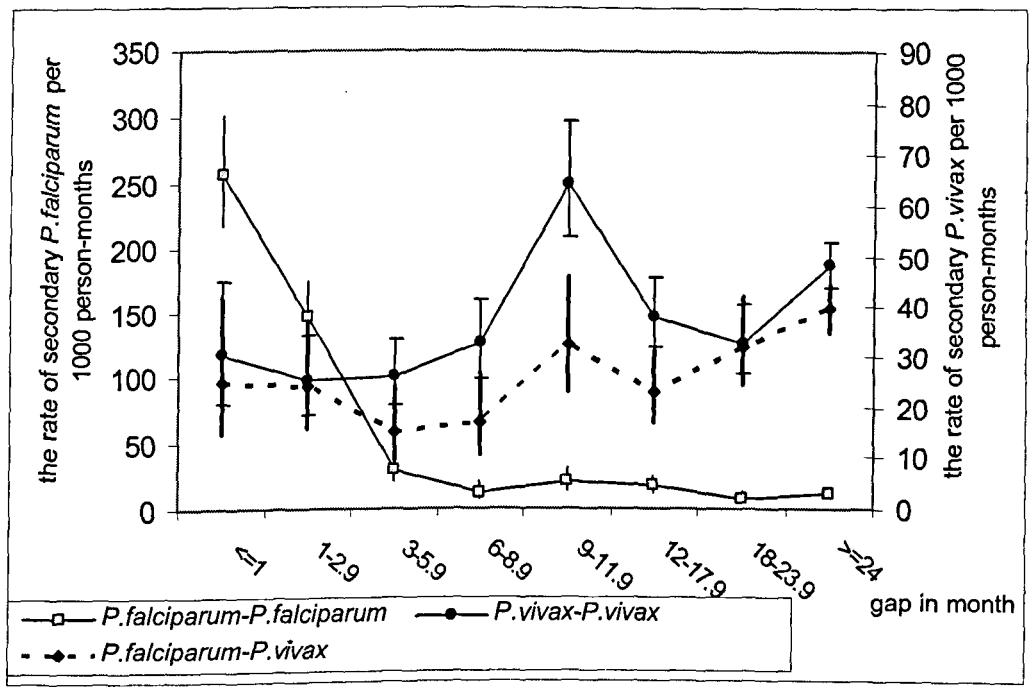

Figure 3-12: The monthly secondary attack rates, classified by the gap and the species in the first and second attacks. *P. falciparum-P. falciparum* rates are shown in left y axis, and the rate in other groups are shown in right y axis

Over the full follow up period, the *P. falciparum* treatment failure rate was around 8.5 times the *P. vivax* treatment failure rate. *P. falciparum* cases usually have higher compliance than *P. vivax* cases, because *P. falciparum* disease is typically more severe than *P. vivax* disease and the *P. falciparum* patients generally seek treatment more actively. Furthermore, the duration of *P. falciparum* treatment is shorter than in *P. vivax* (3 versus 5 to 14 days respectively) *[68]*. Therefore, higher *P. falciparum* treatment failure rate could probably not be explained by lower compliance in *P. falciparum* patients in taking drugs.

The above finding could be explained by a higher frequency of *P. falciparum* resistance to chloroquine. In other words, some of *P. falciparum* cases

might not fully respond to chloroquine, and they contracted the second attack shortly after their first one.

If we assume that *P. falciparum* treatment failure reflects chloroquine resistance, it could be said that *P. falciparum* resistance risk was around 250 per 1000 cases. In other words, a quarter of *P. falciparum* cases showed a second positive blood slide within a month of the first drug dose.

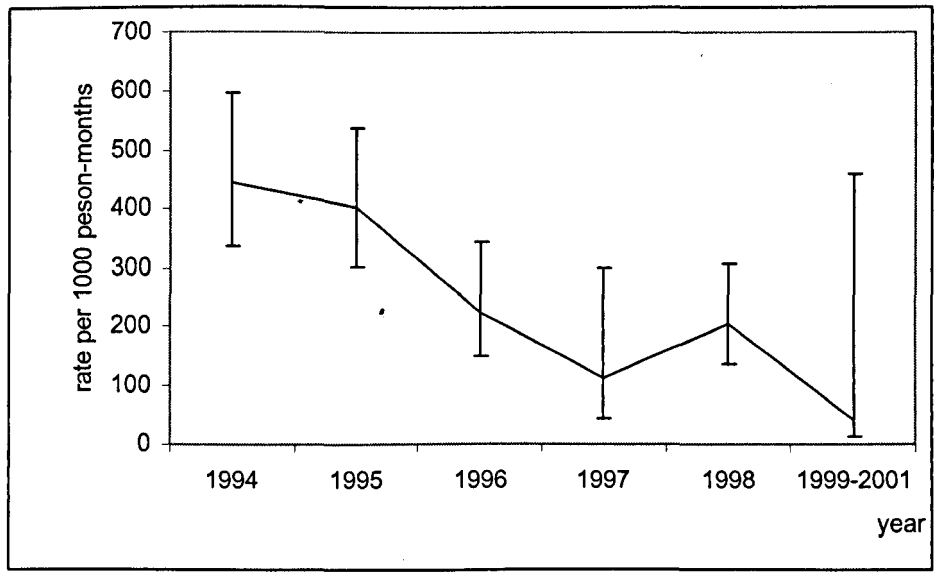

Figure 3-13: The temporal variations of *P. falciparum* therapeutic failure rate

Figure 3-13 illustrates the time trend of *P. falciparum* treatment failure between 1994 and 2001. The rate ratio for year for *P. falciparum* treatment failure was 0.72 (95% CI: 0.65-0.8), meaning that during this period the *P. falciparum* treatment failure rate had a declining trend; the minimum treatment failure rate was less than 100 per 1000 person in 1999-2001.

In contrast to above finding, Edrissian et al. (1989, 1999) show an increasing trend in the frequency of *P. falciparum* resistance to chloroquine in Iran in the last two decades *[62,65]*. However, these data were not based

on *in vivo* tests of clinical response of patients to chloroquine. There is not any precise information about the clinical resistance rate of *P. falciparum* to chloroquine in Iran; most studies estimated this rate based on molecular markers or *in vito* tests *[64-66,115-121]*.

Given the available information, it is very difficult to explain the declining trend of *P. falciparum* treatment failure. It could be due to better health care provision or greater public awareness of malaria which might increase the level of compliance. Furthermore, it should be noted that during this period of time, Iran gradually accepted fewer Afghani immigrants. Although there is not any clear evidence to show that resistant strains came from Afghanistan, local health workers usually acknowledge that Afghani immigrants had more resistant *P. falciparum* infections. Therefore, some of this trend might be due to the changes in immigration patterns.

On the whole, the *P. falciparum* treatment failure was around 250 per 1000 person between 1994 and 2001. WHO recommends changing the first drug of choice if the frequency of resistance is more than 20% *[67,109,110]*. The treatment failure had a declining trend; nevertheless, it seems that *P. falciparum* treatment failure is a very important issue in the management of malaria in Kahnooj, and the health system should be sensitive to this issue. To be proactive and verify the treatment protocol, the health system needs more accurate and up to date information about the chloroquine resistance rate. A precise monitoring system is needed to provide the necessary information.

3.7.3.2. *P. vivax* relapse rate

As discussed in Section 3.7.3, differentiation of re-infection from relapse is very difficult; therefore, most papers reported a combined risk or rate of re-infection and relapse, or estimated the relapse risk based on the frequency of *P. vivax* during the non-transmission season. In this section, a novel method is applied to estimate the relapse rate. The observed secondary *P. vivax* rate in those who had primary *P. vivax* infection is a combined rate of re-infection and relapse. The secondary *P. vivax* attack rate in whose who had primary *P. falciparum* infection could estimate the re-infection rate with *P. vivax*. Therefore, any differences between secondary *P. vivax* attack rates in those who had primary infection with *P. falciparum* and *P. vivax* can be interpreted as the relapse rate.

In this approach there are two assumptions. The first one is that *P. falciparum* induces the same immunity as does *P. vivax* to the acquisition of new *P. vivax* infections. The overall annual infection risk in Kahnooj was very low in recent years (Table 3-1). Due to low exposure risks and a few repeated infections, people did not acquire considerable immunity. Also, the cross immunity between species was not important particularly in low endemic areas (Section 7.7.5).

The other assumption is that the *P. vivax* inoculation rate in *P. falciparum* positive people was the same as in *P. vivax* positive people. Since the temporal distribution of these two species was not similar, the result was adjusted for season. However, the spatial distribution and age group of *P. falciparum-P. vivax* and *P. vivax-P. vivax* groups were comparable ($p>0.13$).

Figure 3-12 compares the rates of secondary attack in those who had primary infection with *P. vivax* and *P. falciparum*, classified by the gap between two attacks. The patterns of changes in both groups were comparable; they had a sharp peak around one year, which reflected the annual variation of *P. vivax*. In other words, irrespective of the previous history of *Plasmodium* spp infections, the risk of *P. vivax* re-infection had an annual cycle. These two lines were very close before 3 and after 18 months gap, i.e., the estimated relapse rate was trivial before 3 and after 18 months. The relapse rates were considerable between 3 and 18 months, however, they were statistically significant only between 6 and 18 months. The monthly relapse rates at 3-5, 6-7, 8-12 and 12-17 months after the initial infection were 10.9, 15.7, 31.9 and 15.1 per 1000 person-months respectively; and the relapse rate in the first 2 years after the primary attack was 1.2 per 1000 person-months. Converting the rate to risk, the relapse risk in one and two years after the primary attack were 16.8% and 24.5% respectively.

The pattern of relapse in Kahnooj was compatible with the relapse pattern in other temperate areas *[104];* i.e., the *P. vivax* relapse rate before 3 months was very low and its maximum rate was observed around one year after the primary attack.

The risk of *P. vivax* relapse was around 25% in two years, which is very close to the relapse risks which has been observed by Rowland and Durrani (1999) in Pakistan in the treatment arm with 14-day primaquine regime, *[100]* and by Prasad et al (1991) *[99]* and Leslie et al. (2004) *[106]*. However,

it is less than reported risks in other studies *[98,102,103,105,122];* but it should be noted that these usually reported the joint risk of re-infection and relapse; but this analysis estimated the relapse risk separately.

Rowland and Durrani (1999) *[100]* showed in his clinical trial that, even with 14-day primaquine, around 34% of *P. vivax* cases show at least one episode of relapse/re-infection. Leslie et al. (2004) showed that around 20% of *P. vivax* cases showed at least one secondary attack within nine months.*[106].* The estimated relapse risk in this study was 25% in the first two years. However, it is an estimation of just relapse. Therefore, from a practical point of view, it seems that the anti-relapse treatment had an acceptable effectiveness in Kahnooj. Although the national protocol in Iran has recommended a 14-day regime of primaquine, general practitioners and health workers have the right to make a final decision about the duration based on the compliance of cases. Unpublished reports from a few locations in Kahnooj show that around 80%-90% of cases receive a 14-day regime.

On the other hand, the relapse risk is high enough to reduce the accuracy of predictive models if they do not adjust their results based on relapse risks. In other words, lower accuracy of *P. vivax* models than *P. falciparum* ones can be expected, if the relapse risks are ignored (Section 3.7.4.6).

3.7.4. Temporal variation

In the following sections, the temporal variations in the numbers of *P. falciparum* and *P. vivax* cases are modelled. Section 3.7.4.1 illustrates the annual variation of malaria, and meteorological variables in Kahnooj

between 1987 and 2001; and assesses their relationships. Section 3.7.4.2 explains the variations in the number of malaria cases between dekads (10 day periods). Sections 3.7.4.3 and 3.7.4.4 assess the seasonality and time trend, and the effect of meteorological factors on malaria respectively. Section 3.7.4.5 presents the results of time series analysis; and the last section (3.7.4.6) discusses the accuracy of the final model in predicting malaria.

The results of this section explore the feasibility of models in prediction of malaria in temporal span.

3.7.4.1. Annual data

To evaluate the correlations between meteorological factors and malaria, annual malaria data for Kahnooj in last 16 years (1987-2003) are analysed in this section. The specific meteorological factors considered are total annual rainfall, the annual mean of daily minimum, maximum and mean temperatures and relative humidity. The malaria situation is described by the following standard indices:

1. Annual Blood Examination Rate (ABER): the number of slides examined for malaria parasites per one hundred population per year
2. Smear Positivity rate (SPR): the percentage of positive slides
3. Annual Parasite Index (API): the number of new malaria cases per one thousand population per year
4. Annual *falciparum* Index (AFI): the number of *P. falciparum* cases per one thousand population per year

Malaria had a wide range of variation during the last 16 years (Figure 3-14). API and AFI rose between 1987 and 1993; the maximum API and AFI were 55.4 and 34.4 per one thousand respectively in 1993. They subsequently decreased at a more or less constant rate. The ratio of maximum to minimum API was 27.7, the corresponding figure for AFI was 191.4; i.e., the risk of malaria in 1993 was around 28 times that of 2002; and the risk of *P. falciparum* in 1993 was around 190 times that of 2001.

Variations in ABER and SPR did not follow those in API. ABER peaked one year later in 1994 and remained at a high level for several years, which might reflect the health system's alertness after the 1993 outbreak. In other words, after the 1993 outbreak, the surveillance system might be more vigorous and more blood slides were taken between 1994 and 1998. Because of high ABER between 1994 and 1998 and decreasing trend of API, the SPR dropped sharply after 1993.

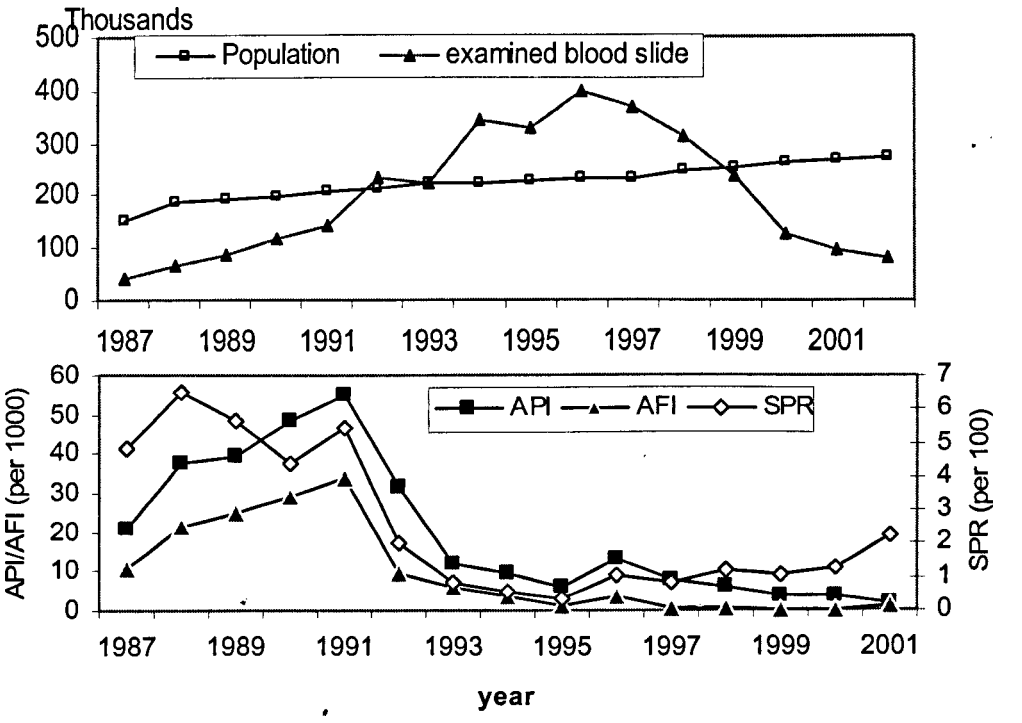

Figure 3-14: The annual malaria indices between 1987 and 2002 in Kahnooj (See API, AFI and SPR definitions in the text)

Different calendar years were used to check the association between rainfall and malaria. Rain comes mostly between November and March in Kahnooj (Figure 3-3); and malaria transmission occurs predominantly between March and October (Figure 3-16). Therefore, any rainfall in the last two months of year does not have any impact on the malaria in that year, but it can influence malaria risk in following year. Therefore, the total amount of rainfall was computed based on the daily data of the last two months from previous year and the first ten months of that year; i.e., from November to October.

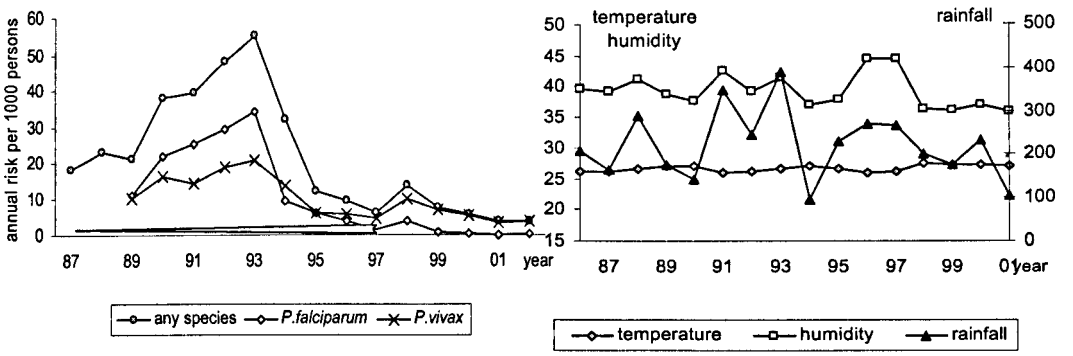

Figure 3-15: The annual malaria risk and meteorological variables between 1987 and 2002 in Kahnooj district, the mean temperature is measured in Celsius, relative humidity in percent and total rainfall in millimetre

Visual examination of Figure 3-15 does not suggest any strong associations between annual meteorological factors and malaria risk. Both species, particularly *P. falciparum* had declining trends, with two peaks in 1993 and 1998, while the relative humidity and temperature did not have considerable variations. However, the high amount of rainfall in 1991-93 might explain the peak of malaria in those years to some extent.

In contrast to minimum temperature and humidity, mean and maximum temperature and humidity had positive associations with malaria (Table 3-7). The Pearson correlation coefficient between the annual malaria risk and humidity was negative; the correlation with minimum temperature was very close to zero. On the other hand, maximum temperature and rainfall had considerable positive associations with annual risks. There is a suggestion that *P. vivax* was more strongly correlated with maximum temperature and *P. falciparum* with rainfall.

Table 3-7: Pearson correlation coefficients between annual risk of malaria and meteorological variables in Kahnooj 1887-2001

Meteorological factor	API[1]	AFI[2]	AVI[3]
Minimum temperature	-0.02	-0.01	-0.04
Maximum temperature	0.40	0.33	0.46
Mean temperature	0.18	0.15	0.19
Humidity	-0.12	-0.09	-0.14
Rainfall	0.45*	0.54*	0.40*

1: Annual parasitic index
2: Annual *P. falciparum* index
3: Annual *P. vivax* index
*p<0.05

Less than one third of annual malaria variation could be explained by meteorological variables. *P. falciparum* had the strongest correlation with rainfall (r=0.54); i.e., rainfall explained around one third of *P. falciparum* variation (R^2=0.292). Adding all meteorological variables into the regression model did not improve the goodness of fit considerably, the adjusted R^2 was 0.31. Therefore, even the best model can not explain two thirds of *P. falciparum* variability. The R^2 for *P. vivax* was 0.23.

In fact, the summaries of annual meteorological variables are not useful predictors. They cannot be used to predict the situation of malaria in the future. Also, the above analyses showed that they predict less than one third of malaria variations. In addition, malaria is transmitted in around 5 months effectively (April to August) in Kahnooj, and climate effect during this period of time could be masked easily by the data in other months. Due to this limitation, in the following sections the models link the summaries of meteorological variables in dekads (every 10 days) to the malaria risk with different lags to analysis their relationships more precisely.

3.7.4.2. Description of dekad data

This section explains the temporal variations of malaria. The number of cases in dekads is illustrated, and is compared with the result of mean-median smoothing. Then the shape of annual *P. vivax* and *P. falciparum* epidemics are compared.

Malaria has a clear seasonal pattern in Kahnooj (Figure 3-16). During the summer (April to September), malaria risk was more than 100 cases per 100,000 population per dekad; however, during the winter (November to February), the risk was less than 10. These seasonal variations were observed for both *P. falciparum* and *P. vivax* species.

In contrast to *P. vivax*, *P. falciparum* had a declining trend; however, both of these species had two prominent peaks in 1994 and 1998. In 1994, the maximum risk of *P. falciparum* was around 100 cases per 100,000 population per dekad which fell to less than 5 in 2001 and 2002. In contrast, the time trend in the annual *P. vivax* risk was not seen visually; maximum risk in 1994 and 1998 were around 100 cases per 100,000 populations per 10 days, which was around twice of the risks in other years. The *P. falciparum* peak in 1994 was around three times of that in 1998; however, the *P. vivax* peaks in 1994 and 1998 were comparable.

The shape of smoothed curves in *P. falciparum* and *P. vivax* differ significantly. Using local mean-median smoothing method with a span of three dekads, the coarse variations between consecutive dekad risks were removed. A span of three dekad was chosen because one month captured the highest amount of variability. The *P. falciparum* curve had a prominent peak at the end of each summer, except in 1997 which had a short bimodal

curve. However, *P. vivax* showed bimodal curves in summers except in 1998 which had a very sharp peak at the end of summer; the first *P. vivax* annual peaks were around April-May, and the second ones were around August-September.

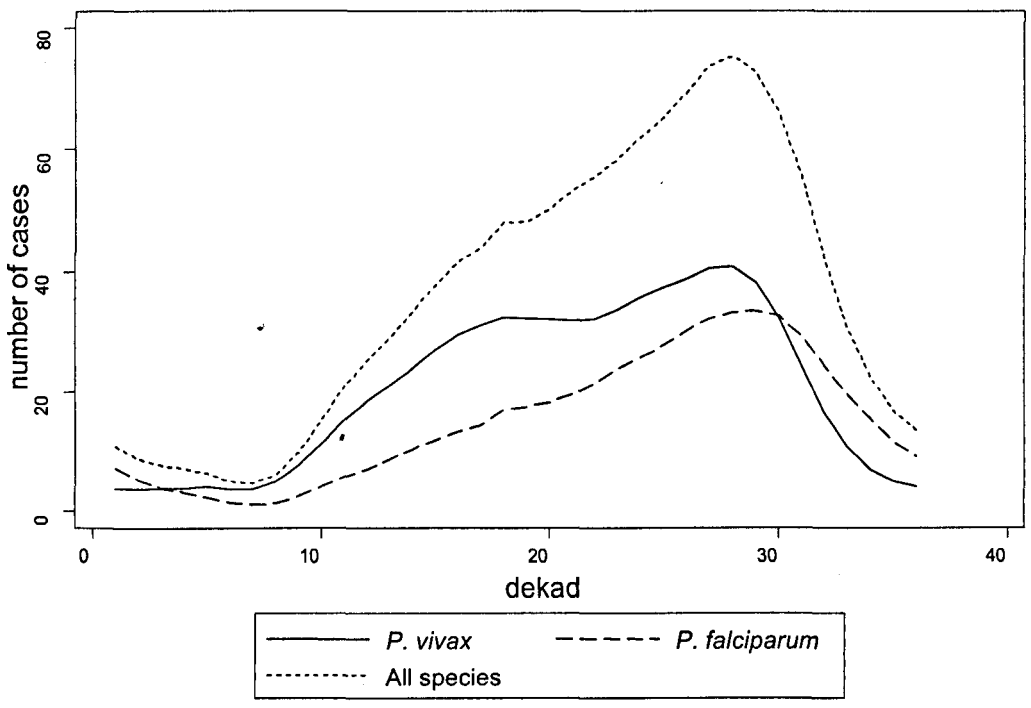

Figure 3-16: Temporal variations of malaria over a year; the observed numbers classified by species

A bimodal *P. vivax* curve is frequently reported in temperate climates *[37,123,124]*. Generally the early peaks are explained by *P. vivax* relapse and the late peaks by high transmission. Bouma et al. (1996) *[37]* added that low humidity in May and June in his study field in Pakistan might decrease the life span of vectors, which caused a drop in the number of cases between two *P. vivax* peaks.

It is not easy to explain the early *P. vivax* peak by relapse based on the observations in this study. There were not any obvious associations between the late *P. vivax* peaks of one year and the early peak of the following year. For example, the high incidences of *P. vivax* later in the summers in 1994 and 1998 were not followed by a sharp peak at the beginning of 1995 and 1999.

In addition, the *P. vivax* relapse rate was low in Kahnooj. Section 3.7.3.2 shows that the maximum relapse rate was observed around 8-12 months after the primary attack, and was 32 per 1000 *P. vivax* cases per month. Therefore, it would be hard to expect a noticeable peak just due to relapse in the beginning of summer.

However, *P. vivax* relapse could play an important role in the transmission chain between years. In the beginning of each summer, mosquito density increases and even small number of *P. vivax* relapse cases could be enough to infect the vectors, and start new transmission cycles.

An alternative explanation for the bimodal curve is unsuitable climatic conditions in the mid-summer *[37]*. The effect of very hot temperature and low humidity in the mid summer is assessed in the following sections.

3.7.4.3. Seasonality and time trend

This section models the seasonality and time trend of malaria. Applying a Poisson model adjusted for population, the optimum sinusoidal model is

used to explain seasonality. The time trend is modelled by the linear and quadratic effects of year.

Seasonality is a well known phenomenon in malaria. Gill (1938) *[123]* explained the different seasonal patterns in temperate, sub-temperate, tropical and equatorial zones. Macdonald (1953) *[124]* used a mathematical approach to assess the malaria epidemic curve. Since then, a great deal of research has been conducted to explain the effect of climate on the seasonal variations of malaria.

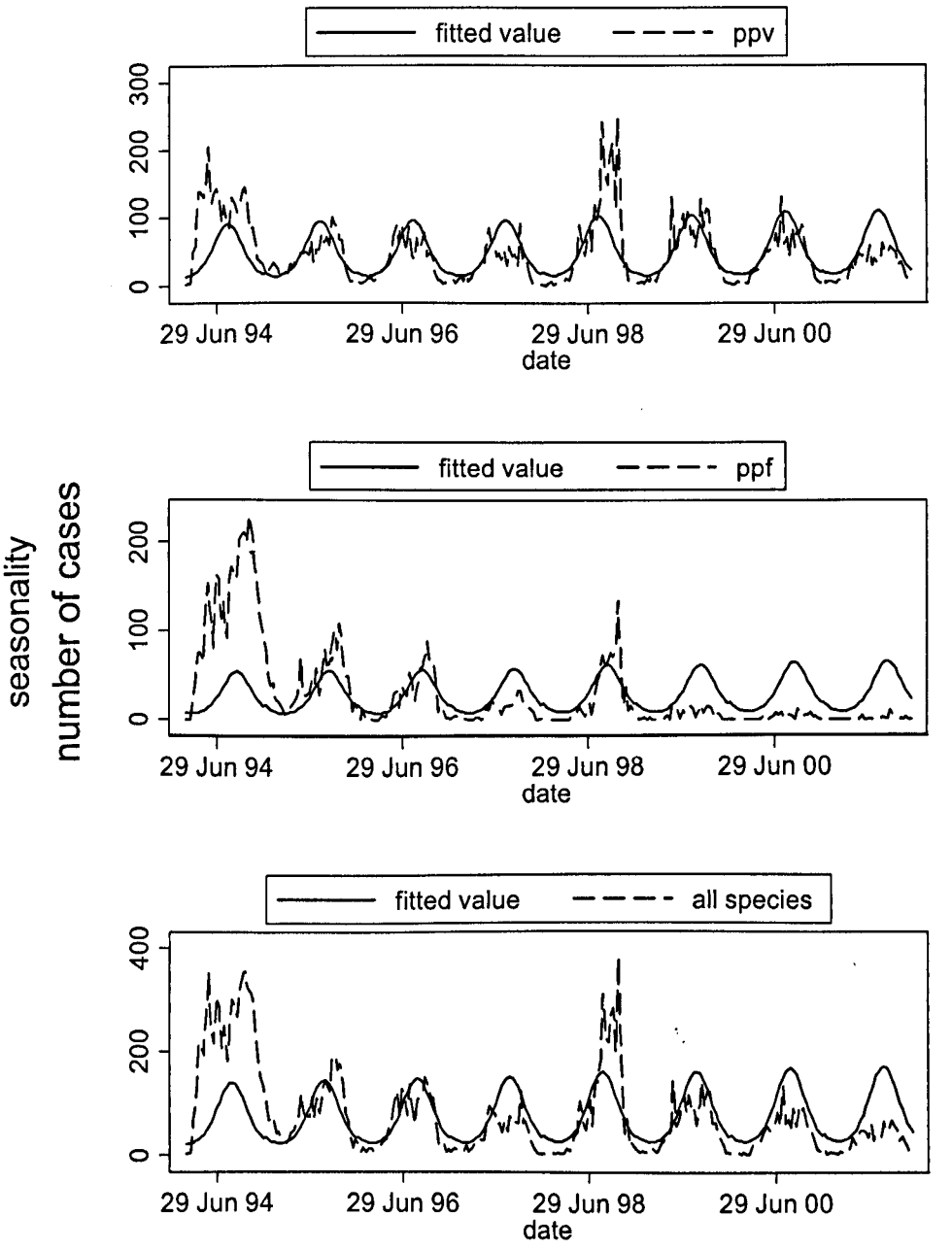

Figure 3-17: The seasonality of malaria classified by species, the observed numbers (dashes) and model estimated number (solid line) in the 'fitting' part of the dataset

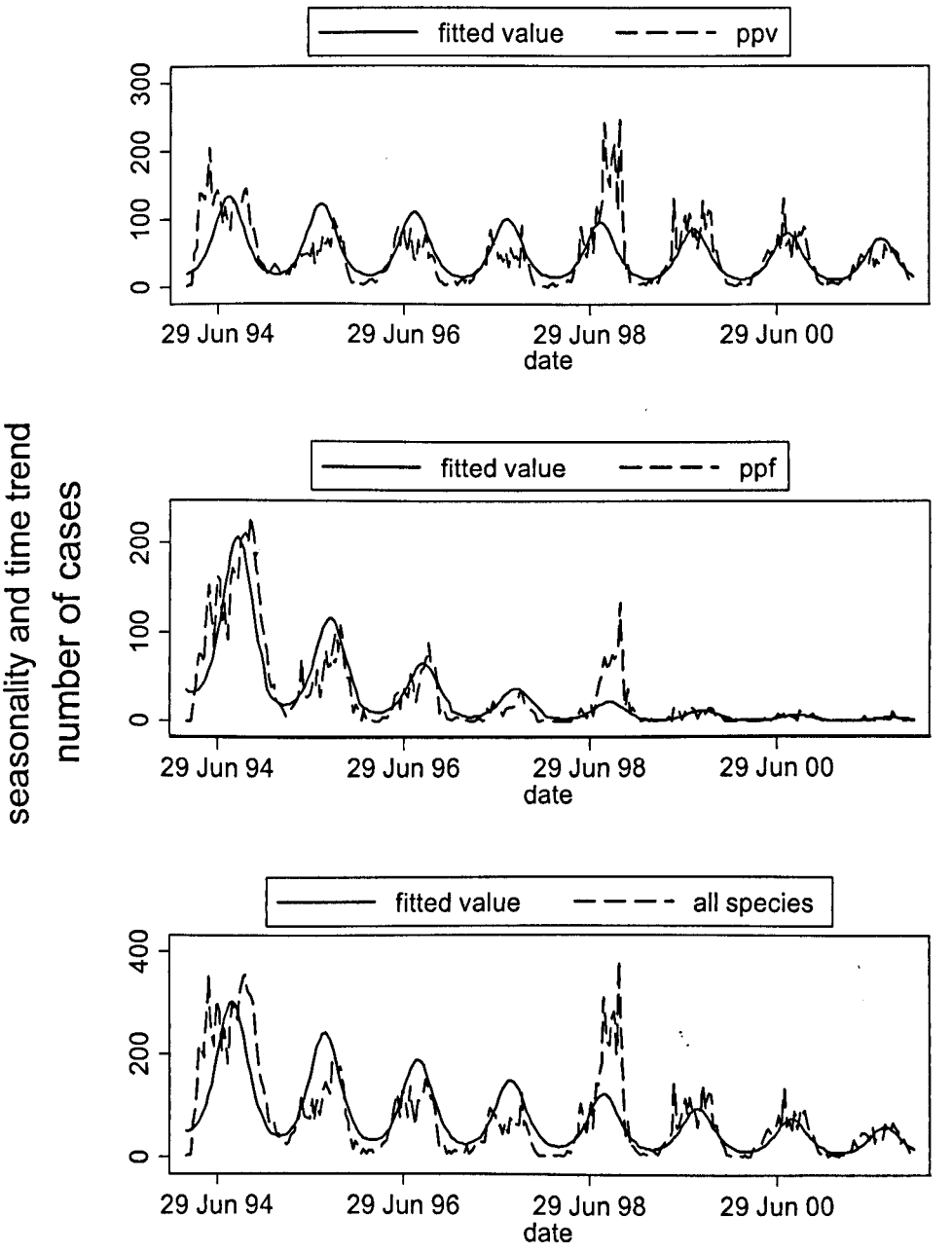

Figure 3-18: The seasonality and time trend of malaria classified by species, the observed numbers (dashes) and model estimated number (solid line) in the 'fitting' part of the dataset

Figure 3-17 shows clear seasonal pattern in malaria in Kahnooj particularly in *P. vivax*, which was modelled sinusoidally (Section 3.4.7). LR tests

showed highly significant p-values (<0.0001) in *P. vivax*, *P. falciparum* and all species. The pseudo R^2 of *P. vivax*, *P. falciparum* and all species were 0.42, 0.18 and 0.32 respectively.

The phases (ε) of these models imply that *P. vivax* peaked around one month after *P. falciparum*. The fitted peaks of *P. falciparum* and *P. vivax* were on 12th of July and 16th of August respectively.

Nevertheless, prediction of the peaks of species incidences annual based on the peak of the sinusoidal curves might mask the actual annual variations. Section 3.7.4.3 explains the differences between *P. falciparum* and *P. vivax* curves; *P. vivax* had bimodal annual variations with two peaks in early and late summer; however, *P. falciparum* had one annual peak mostly in late summer. Therefore, other factors such as meteorological factors might change the peak of actual epidemic curves.

The *P. falciparum* and all species graphs had clear decreasing time trends. *P. falciparum* risk in summer 1994 was around 200 times that risk in summer 2002. In contrast, the *P. vivax* graph does not show any clear time trend. Nevertheless, the linear effect of year in all models were significant (LR test: $p<0.001$).

Seasonality and time trend explain *P. falciparum* variations better than *P. vivax* variations (Table 3-8). Having adjusted for seasonality and linear effect of year, the pseudo R^2 for *P. vivax* was 0.49, the corresponding R^2 for *P. falciparum* and all species were 0.76 and 0.6 respectively (Table 3-8). Nevertheless, the quadratic effects of year did not improve R^2 values (all p-values were greater than 0.15).

In summary, it could be concluded that the sinusoidal model with linear time trend is a simple and appropriate method to explain some part of temporal variations, particularly *P. falciparum* variations. The clear seasonal pattern of malaria in Kahnooj and its time trend could be modelled on sine transformation of time and linear effect of year.

3.7.4.4. Climate effect

This section examines the effect of rainfall, temperature and relative humidity on the variation of malaria classified by species. First, the annual variation of mean monthly temperature and relative humidity is illustrated. Then, the importance of meteorological variables is assessed, incorporates different lag effects. Finally, the optimum temperature and relative humidity for transmission of species are estimated.

The meteorological variables ranged widely in Kahnooj between 1994 and 2001. Temperature ranged between -1 and 50°C; the minimum and maximum recorded mean daily temperatures were 5 and 42°C respectively. In 50% of days the mean daily temperatures were more than 28°C, and in 20% of days more than 36°C. Relative humidity varied between 5 and 99%. The mean daily relative humidity was less than 23% and 36% in fifty and eighty percents of days respectively; and in only ten percent of days it was more than 62.4%. The minimum and maximum annual rainfalls were 93 mm in 1994 and 371 mm in 1996 respectively.

The temporal variation of temperature and relative humidity had opposite patterns (Figure 3-18). Temperatures peaked between June and September, when relative humidity was lowest. In contrast, maximum relative humidity and minimum temperature were recorded in January and February. The Pearson correlation coefficient between mean daily temperature and relative humidity was -0.53 (95% CI: -0.53, -0.56). This correlation coefficient was -0.52 (95% CI: -0.5,-0.55) in transmission months, i.e., April to September.

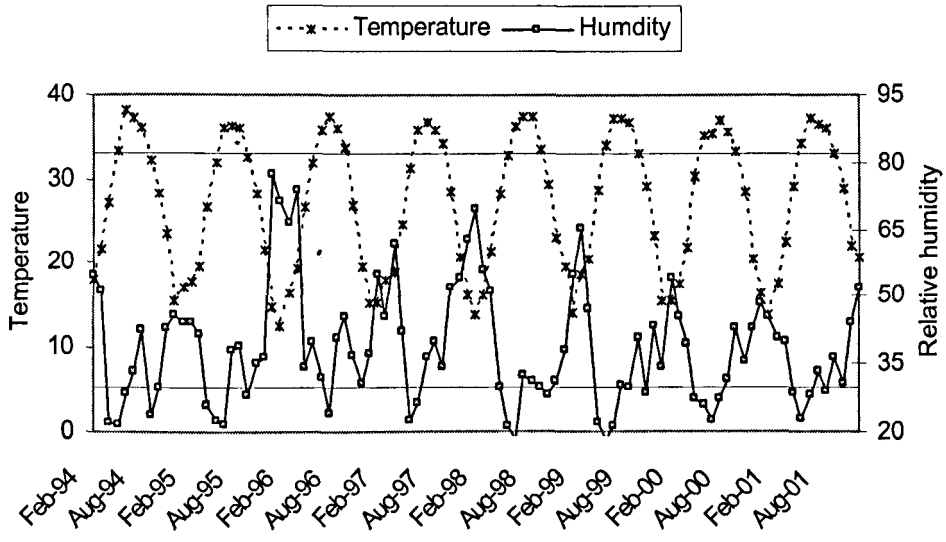

Figure 3-19: The annual variation of mean monthly temperature and relative humidity. The two horizontal lines cross the vertical axes at 33°C and 30% relative humidity; generally, the optimum condition for malaria transmission is when both curves are above these two lines. The cut off points are estimated in following paragraphs

A two to four dekad lag maximizes the correlation coefficient between the number of cases and meteorological variables. The associations were assessed based on the number of cases in each dekad and the mean temperature and relative humidity in the preceding 1-6 dekads. The maximum associations were detected between *P. falciparum* and meteorological variables with a three dekad gap, while maximum

associations were observed between *P. vivax* and meteorological factors with two dekad gap. (Also see following paragraphs about the optimum lag between climate and species.)

Associations between the number of cases and meteorological factors with 0-6 dekad lags were very close. R^2 values between *P. vivax* and mean temperature, for 0-6 lags, ranged from 0.37 to 0.4; and with humidity ranged from 0.19 to 0.21. The corresponding ranges between *P. falciparum* and mean temperature, and humidity were 0.04-0.05 and 0.03-0.04 respectively.

P. vivax had stronger positive association with temperature than *P. falciparum*; but the relationships between both species and humidity were not strong. Having applied the partial correlation coefficient, the effects of temperature and humidity were adjusted for each other. The correlation coefficients between *P. vivax* and temperature, and humidity with a two dekad lag were 0.48 ($p<0.0001$) and -0.08 ($p=0.19$) respectively. The corresponding coefficients in *P. falciparum* with a three dekad lag were 0.12 ($p=0.048$) and -0.07 ($p=0.21$).

Temperature and relative humidity explained *P. vivax* variations better than *P. falciparum* variations. The pseudo R^2 between *P. vivax* and mean temperature and relative humidity with two dekad lag was 0.4. However, the pseudo R^2 values between *P. falciparum* and all species with mean temperature and relative humidity were 0.06 and 0.24 respectively.

Rainfall was defined as the total amount of rain between the previous November and the preceding two dekads. Section 3.7.4.1 explained that the importance of rainfall in one rainy season on the malaria in the following

transmission season. Therefore, the rainfall was measured from previous November. Although the amount of rainfall is negligible in the transmission season, to keep the rainfall data compatible with other meteorological variables, it was included with lags of up to two dekads.

On top of the seasonality and trend effects, meteorological factors still explained a significant amount of malaria variation, particularly in *P. vivax*. Reviewing the pseudo R^2 in Table 3-8 shows that temperature and relative humidity in the previous two dekads could explain the malaria variations as well as temperature and relative humidity in all 6 previous dekads. Although the quadratic effect of temperature and relative humidity were significant in both *P. vivax* and *P. falciparum* models ($p<0.0001$), the changes in pseudo R^2 in the *P. vivax* model were much more than in the *P. falciparum* model (6% versus less than 1%). Adding the optimum combinations of temperature, relative humidity and annual rainfall to the seasonality and time trend improved the pseudo R^2 by around 6% and 17% in the *P. falciparum* and *P. vivax* models respectively.

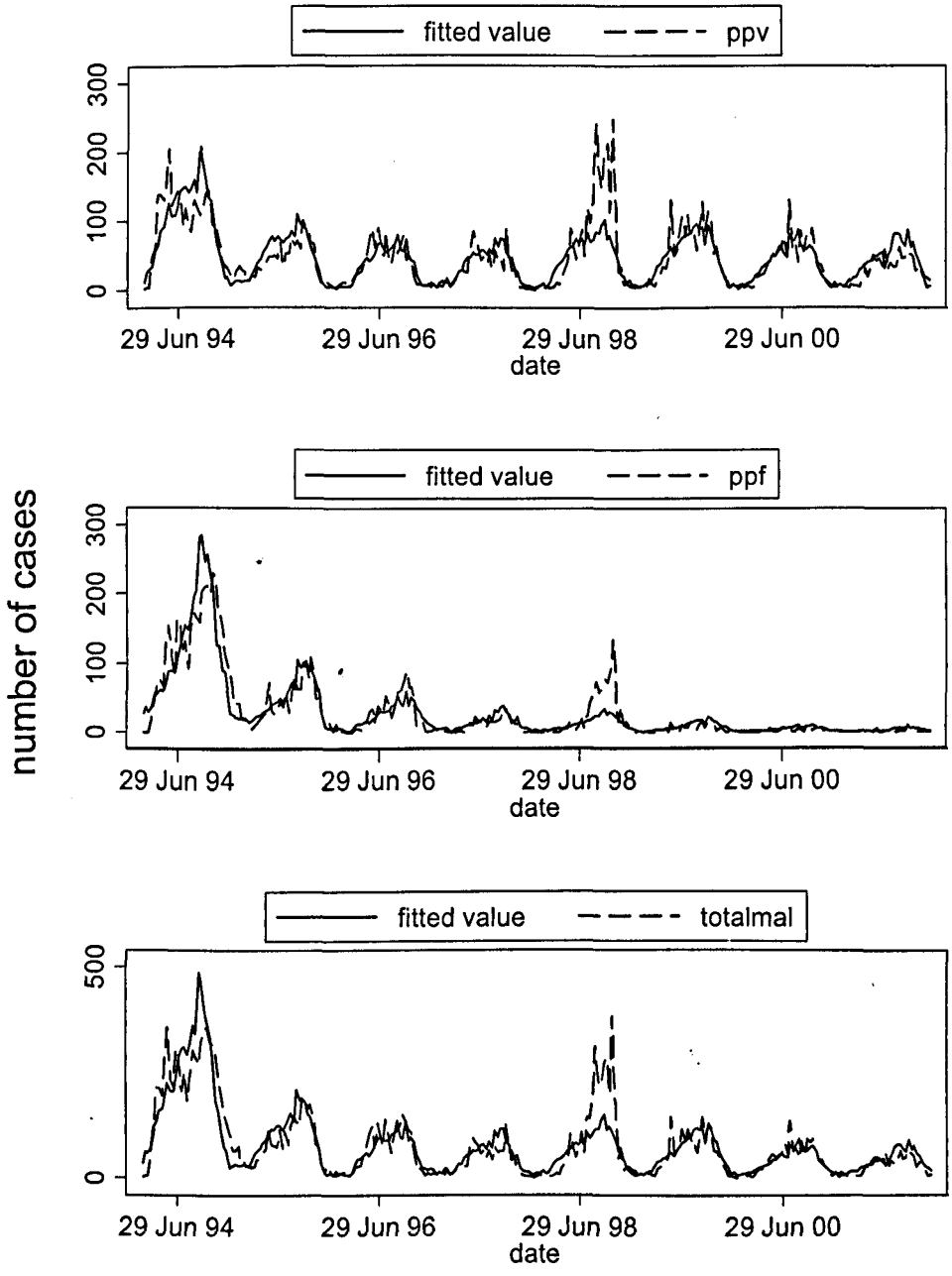

Figure 3-20: The fitted values of models based on seasonality, time trend and meteorological variables classified by species, observed numbers (dashes) and model estimated number (solid line)

The model based on seasonality, time trend and meteorological factors explained *P. falciparum* variations better than *P. vivax*. The pseudo R^2 in the *P. falciparum* model was 0.8; the corresponding value for *P. vivax* was 0.66. Figure 3-19 illustrates that except for the epidemic in 1998, the models explained the variations of species quite precisely.

Lower accuracy of the *P. vivax* model could be due to relapse. As explained in Section 3.7.3.2, generally *P. vivax* relapse is common even with a complete dose of an anti-relapse drug. In Kahnooj, around 25% of *P. vivax* cases showed relapse during the first 2 years after the primary attack. To address the relapse effect on the goodness of fit of these models, the autocorrelation between the numbers of cases will be taken into account in Section 3.7.4.5.

The optimum meteorological conditions were estimated by solving the differential Poisson quadratic equation of lagged mean temperature and relative humidity. In the models, temperature and relative humidity and their square terms were entered. The solving first differential equation in respect to either of the variables estimates the turning point of the graph; i.e., the value of temperature or relative humidity which the disease risks were minimum or maximum.

Compared to *P. falciparum*, *P. vivax* peaked during higher temperature and lower relative humidity. The optimum mean temperatures for *P. vivax* and *P. falciparum* were 35 and 31.1°C respectively. The corresponding values for relative humidity were 27.3 and 32%. The optimum temperature and relative humidity for all species were 34°C and 30.1 percent.

The above estimated optimum temperatures were slightly greater than the estimated temperatures of other laboratory studies that assessed the effect of temperature on mosquito development *[13]*. These differences might be due to the fact that the rate of mosquito development is an important factor in malaria transmission; however, other factors such as parasite development inside the mosquito and the mosquito survival rate are also important. In addition, it should be mentioned that in the field, mosquitoes reside in the optimal microclimate *[38]*.

According to Figure 3-18, the optimum climate for malaria transmission was observed generally at the end of summer around August and September. During these two months, temperature and relative humidity were above 33°C and 30% respectively

The local minimum in the *P. vivax* curve in mid summer (Figure 3-16) could be explained by very hot weather. The mean of temperature in June and July were 36.5 and 37.4°C respectively, which were much higher than the estimated optimum temperature (35°C), while the mean temperatures before and after mid summer in May and August were 33.1 and 35.4°C respectively. The bimodal annual curve was more obvious in 1994, 1999 and 2001. The maximum mean temperature in mid summer in these three years were 38, 37.4 and 37.2°C respectively, which were higher than the mid summer temperature in other years ($p=0.02$). However, there was no clear association between mid summer relative humidity and the drop in *P. vivax* risk. It should be mentioned that the spatial variation of relative humidity in summer was wider than in temperature. Therefore, to have a clearer view about the effect of relative humidity in mid summer drop, spatial analysis should be applied (Section 3.7.5).

Low humidity in the beginning of summer might repress the early *P. falciparum* peaks because *P. falciparum* was more sensitive. In contrast to *P. vivax*, *P. falciparum* did not show a bimodal curve in most years. More precisely, it could be said that the early peaks of *P. falciparum* might be inhibited. Therefore, *P. falciparum* might not find appropriate conditions in most years to show early peaks similar to *P. vivax* peaks.

3.7.4.5. Time series analysis

This section analyses the autocorrelation in disease risk in the consecutive dekads. Plausible mechanisms for autocorrelation are explained, and its magnitude assessed. Then, the goodness of fit for models, adjusted for autocorrelations are presented.

Dekadal risks of disease are dependent on each other. Malaria depends on climate; therefore, much of its autocorrelation might be explained by autocorrelation in meteorological variables. Malaria as an infectious disease has a transmission cycle between human, mosquito and human. Therefore, the infection load in each part of this circle can be passed to the other parts over time. This dependency can be explained by the following mechanisms:

1. Frequency of infectious vectors: infectious vectors can survive for more than one dekad, and influence the number of cases in consecutive dekads. However, since mosquitoes have short life spans (a few weeks) in the transmission season, it seems that this factor causes autocorrelation only over lags of one or two dekads.

2. Frequency of infectious people: malaria transmits from infectious subject to healthy people via the vector. Therefore, a high number of infectious subjects in one dekad can be an indicator of another peak in the following dekads. The gap between two peaks, i.e., the duration of the transmission cycle, is determined mostly by climate. On average, it is expected to be two to four dekads in the transmission seasons.

3. Control programs: the health system usually enhances its activities after an epidemic; also, people may be more aware of malaria and use more preventive methods. Therefore, a drop may be observed after a prominent epidemic. This would induce a negative autocorrelation rather than the positive ones expected from the previous two mechanisms.

4. Relapse: in contrast to the other mechanisms, *P. vivax* relapse may generate positive autocorrelation in risks with large lags; i.e., months or even years.

Having adjusted for seasonality, time trend and climate, the autocorrelation patterns in *P. falciparum* and *P. vivax* were comparable. Figure 3-20 shows the correlation and partial correlation coefficients between dekad residual risks with different lags. The maximum autocorrelation coefficients were around 0.6 in both species with a one dekad lag, which shows the importance of autocorrelation. The correlation coefficients between dekad risks with five or less dekad lags were significant; however, the partial

correlation coefficients were significant only with three or less dekad lags in *P. vivax*.

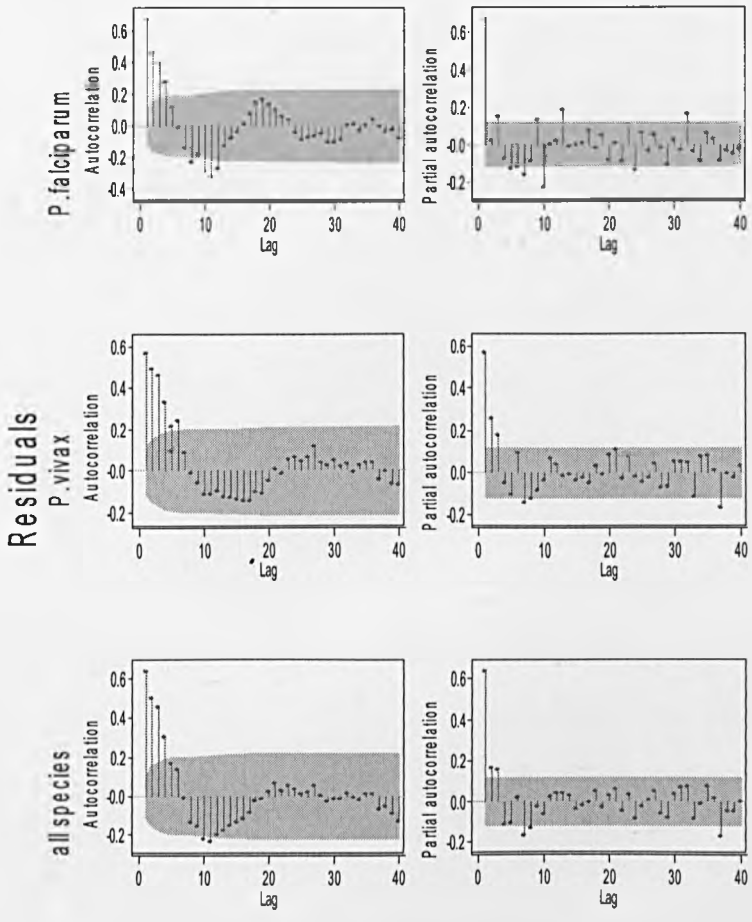

Figure 3-21: The Autocorrelations and partial autocorrelations between the residuals of models, which estimated risks, based on climate, seasonality and time trend; the shaded areas show 95% confidence interval.

Relapses did not have much effect on the autocorrelation pattern in *P. vivax*. The autocorrelation pattern of *P. vivax* risk had a seasonal pattern. Having removed the seasonality and climate effects, none of the correlation and partial correlation coefficients, except the partial correlation

coefficient in dekad 37, were significant with more than 5 dekad lags while, the maximum relapse risks were expected to be observed between 18 and 48 dekad lags (Section 3.7.3.2), the majority of correlation, and partial coefficients ranged between -0.1 and 0.05, and -0.05 and 0.05 after 18 dekad lags.

To minimize the numbers of model parameters, and make it more appropriate for practical use in the field, the number of cases in dekads were aggregated. It is important to check the long term history of malaria in models in order to assess the impact of relapse. Nevertheless, the model would have had numerous parameters, if all dekad data had entered separately. Also, from a practical point of view, long term dekad data are usually inaccessible. To deal with these two problems, sum variables were computed which contain the sum of cases in dekads in the past as follow: last dekad, 2-3 dekads, 4-16 dekads, 17-24 dekads, 25-36 dekads and 37-48 dekads.

It was not possible to compute the sum variables in the first records, because of inaccessibility of past dekadal data. Because of the different number of missing values in different models, their goodness of fit were not comparable. In order to solve this problem and utilize as much of data as possible, missing values were estimated based on seasonality, time trend and meteorological variables.

Having entered the sum variables in the models, the pseudo R^2 improved for *P. vivax* more than of *P. falciparum*. The likelihood ratio test showed significant effects of sum variables in *P. vivax*, *P. falciparum* and both species ($p<0.0001$). The difference between pseudo R^2 values with and

without these sum variables in *P. falciparum* was 0.03 (0.8 versus 0.83), the corresponding differences in *P. vivax* and both species were 0.14 (0.66 versus 0.8) and 0.09 (0.75 versus 0.84) respectively. Having checked the effects of these sum variables separately, the number of cases in the last dekad played the main role; the pseudo R^2 values with just the number of cases in last dekad were numerically very close to their correspondences with the number of cases in all periods; however, the pseudo R^2 with the number of cases in past the two to three dekads were close only in *P. falciparum* (Table 3-8).

It seems that past history of malaria can improve a model's goodness of fit even after taking into account seasonality, time trends and climate effects. Figure 3-20 shows substantial autocorrelations between malaria risks; particularly with a very short lag of a few dekads. While these autocorrelation effects were significant in all models, the changes they induced in pseudo R^2 indicated that they are more important for *P. vivax* than *P. falciparum* (differences between R^2 values in model 15 versus model 16 in following tables).

Table 3-8: The pseudo R^2 of Poisson models classified by the species based on the whole district data; model 18 is the final model

Model number and Explanatory variables		Pseudo R^2		
		P. falciparum	*P. vivax*	All species
M1	Sine transform of time	0.2	0.43	0.35
M2	M1 & linear effect of year	0.76	0.49	0.6
M3	M1 & quadratic effect of year	0.76	0.49	0.61
M4	M2 & mean daily min temperatures in last 6 dekads[1]	0.76	0.5	0.61
M5	M2 & mean daily max temperatures in last 6 dekads[1]	0.76	0.53	0.62
M6	M2 & mean daily mean temperatures in last 6 dekads[1]	0.76	0.51	0.62
M7	M2 & mean daily relative humidity in last 6 dekads[1]	0.78	0.56	0.67
M8	M2 & mean daily min temperatures in last 2 dekads[1]	0.76	0.49	0.61
M9	M2 & mean daily max temperatures in last 2 dekads[1]	0.76	0.52	0.62
M10	M2 & mean daily mean temperatures in last 2 dekads[1]	0.76	0.51	0.61
M11	M2 & mean daily relative humidity in last 2 dekads[1]	0.78	0.55	0.66
M12	M8 & M9 & M10 & M11	0.78	0.55	0.66
M13	M8 & M9 & M11	0.78	0.55	0.66
M14	M13 & rainfall[2]	0.8	0.6	0.72
M15	M14 & quadratic effect of min[1], max[2] of temperature and humidity in last 2 dekads	0.8	0.66	0.75
M16	M15 and the sum of cases in last dekad, and periods with these dekad lags:2-4, 5-16, 17-24, 25-36 and 37-48	0.83	0.8	0.84
M17	M15 and the sum of cases in last dekad	0.83	0.79	0.84
M18	M15 and the sum of cases in 2-4 dekad ago	0.82	0.75	0.79

1: The mean of daily minimum, maximum or mean variable in one dekad (10 days)
2: The total amount of rainfall between last November and 2 dekad ago

3.7.4.6. Final predictive model

This section examines the accuracy of final model. Variables in the final models were selected based not only on their statistical significances, but also their accessibility in real situations in the field. To check the model accuracy, dekadal data were randomly divided into modelling and checking parts as described below. The parameters were estimated based on the modelling part. Differences between observed and fitted values in the checking part were then assessed (Section 3.4.8).

Table 3-8 shows pseudo R^2 values for various models described previously, from the simplest with only the effect of seasonality, up to the most complicated ones incorporates seasonality, time trend, climate effect, and temporal autocorrelation. According to the likelihood ratio test, seasonality and linear effect of time trend were significant ($p<0.0001$). Also, the minimum and maximum of meteorological factors were more important predictors than just their means. Having assessed the impact of meteorological factors using different lags; and based on practical issues, meteorological variables with only two dekad lags show very close pseudo R^2 values to the complex model pseudo R^2 with all data in last 6 dekads.

Although the numbers of cases in the last dekad were more important than the numbers in other dekads, in practice, at least a two dekad lag is needed for a warning system; because health system responses usually experience some delay. The last two rows in Table 3-8 show the pseudo R^2 with the number of cases in last dekad (one dekad lag) or in past two to four dekads (two to four dekad lag). The two R^2 values were close in the case of *P. falciparum* (0.83 versus 0.82); there was a small but noticeable difference

in pseudo R^2 values for the *P. vivax* models (0.79 versus 0.75). Nevertheless, this study aims to check the feasibility of an early warning system. It means that the models should predict the situation of malaria in future; therefore, a 20-day gap is the minimum required gap in practice, to allow time to act. Hence, the final model predicts the number of cases based on the number of cases in two to four dekads ago (model number 18 in Table 3-8).

A quarter of case records were selected randomly to check the accuracy of models (Section 3.4.8).

Parameters were estimated using the final model (number 18 Table 3-8) on the modelling data, and the predicted values and observed numbers compared in the checking part. The accuracy of models was assessed based on the differences between observed and predicted values.

The percentages of under and over predictions were adjusted by transmission period and year. According to the climate and the duration of malaria transmission, four periods were defined: no-transmission: (December-March), early transmission (April-May), mid transmission (June-July), and late transmission (August-November). Then, the over and under predictions of models were divided by the sums of observed cases in each period and year. The overall percentages of over and under predictions were computed as the weighted averages of their corresponding periods; the numbers of observed cases in each period-year were used as the weight.

The over and under predictions of models were all less than 20%, and the maximum errors were seen during the no-transmission season (Table 3-9). The over and under predictions in the modelling and checking parts were very close. The errors were between 30 to 40% in no-transmission period in

both species, while they were less than 20% in mid and late transmission periods. In other words, the models had more accuracy in high transmission periods. Also, the over and under predictions of *P. falciparum* and *P. vivax* models were comparable. In terms of numbers of cases, the maximum error was less than 500 cases per year: around 300 *P. vivax* cases and 200 *P. falciparum* cases.

Nevertheless, the models did not predict the 1998 epidemic. Figure 3-21 illustrates the observed and residual values in the checking part of dataset. The large values of residuals in 1998, particularly in the early and mid transmission period, shows that the epidemic was due to some factors beyond the effect of explanatory variables in these models.

On the whole, it seems that the model based on the combination of predictive variables is an appropriate early warning method (Section 3.8.4). None of the variables by itself had a pseudo R^2 value more than 0.43. However, the combination of variables improved the R^2 to around 0.75. Moreover, all the models predict the risk of disease 20 days in advance; i.e., the gap is enough to warn the health system.

This model does not take into account spatial variations. The above analyses explain the temporal variation of malaria in the whole Kahnooj district. However, Kahnooj is a vast district with considerable variation in climate and environment factors (Section 3.2.2). Therefore, exploring the spatial distribution of malaria is as important as the temporal variation, which is the focus of in the following sections.

Table 3-9: Over and under estimations of final model classified by transmission period, the numbers show the sum of differences between observed and predicted values.

Transmission period		Modelling part[1]		Checking part[2]	
		Over estimation (%[3])	Under estimation (%[3])	Over estimation (%[3])	Under estimation (%[3])
P. falciparum	no[4]	197(24.5)	197(24.5)	64(32.5)	71(36)
	early[4]	147(19)	275(35.5)	35(31.7)	33(29.5)
	mid[4]	156(15.4)	127(12.6)	86(20.4)	46(10.8)
	late[4]	613(17.1)	544(15.2)	135(10.9)	246(19.8)
	whole year	1,113(18.4)	1,113(18.4)	321(16.3)	396(20.1)
P. vivax	no[4]	242(37.7)	165(25.7)	75(38.6)	59(30.4)
	early[4]	231(11.8)	415(21.3)	56(17.2)	79(24.3)
	mid[4]	317(13.1)	296(12.2)	203(20.7)	101(10.3)
	late[4]	664(13)	577(11.3)	274(15.2)	326(18.1)
	whole year	1,454(14.4)	1,454(14.4)	608(18.4)	565(17.1)
All species	no[4]	451(34.2)	345(26.1)	132(33.6)	129(32.7)
	early[4]	338(2.1)	709(4.4)	82(18.9)	122(27.9)
	mid[4]	444(13)	327(9.6)	268(19.1)	121(8.6)
	late[4]	1,093(12.6)	945(10.9)	388(12.8)	509(16.8)
	whole year	2,326(14.4)	2,326(14.4)	870(16.5)	881(16.7)

1: The model was built based on three-quarters of dekad data
2: The fitted value was computed based on the estimated parameters in modelling data
3: Total numbers of over or under prediction divided by total number of cases adjusted for year-period (Section 3.7.4.6 for more details)
4: The average of monthly number of over or under predictions

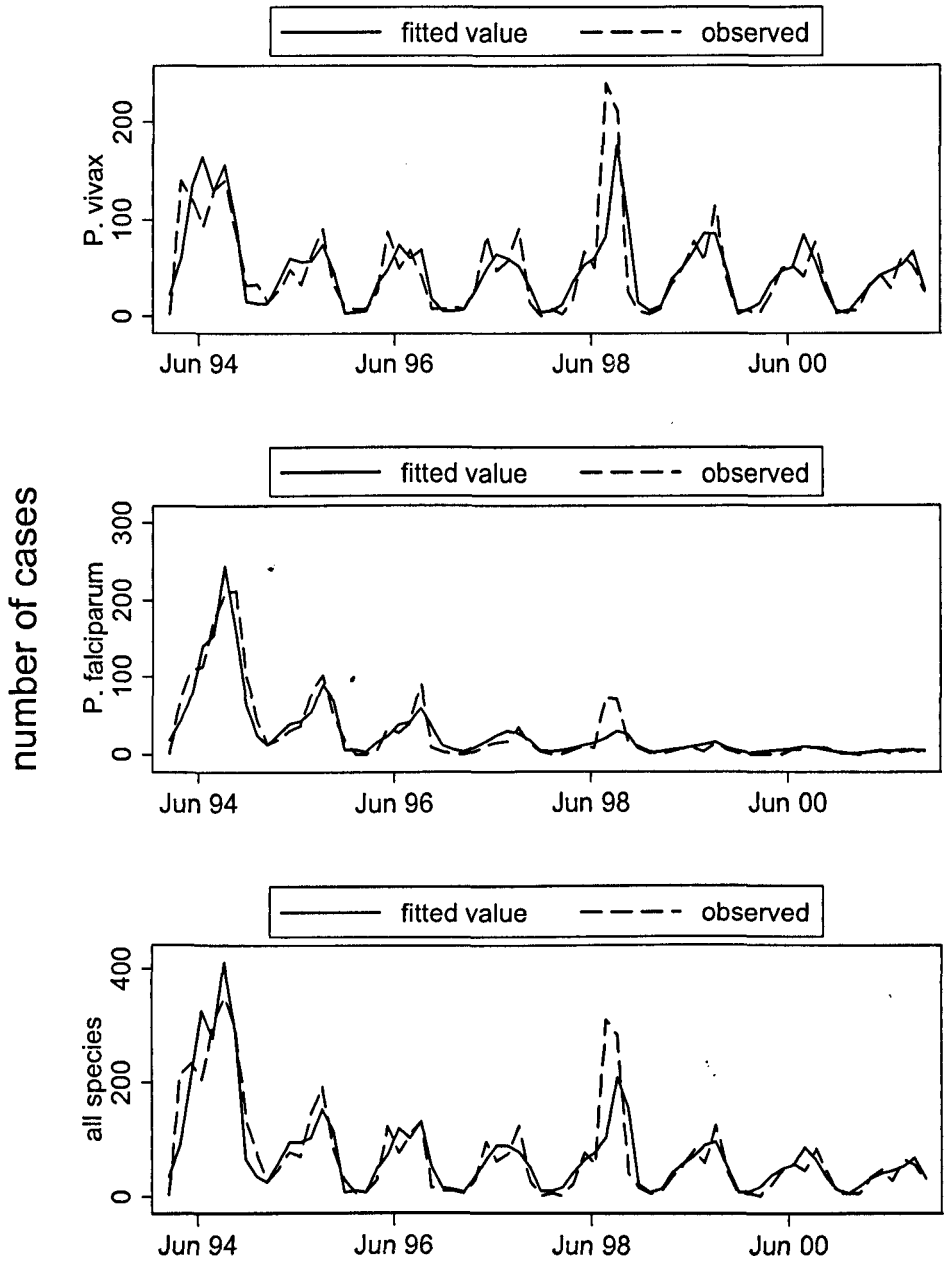

Figure 3-22: The observed data and the fitted value of final model (based on seasonality, time trend, temporal autocorrelation, and climate variables) in the checking part of data

3.7.5. Spatial-temporal variations

These analyses model the spatial and temporal distribution of malaria. The models link ground and remote sensing data with malaria risk. Section 3.7.5.1 explains the geographical distribution of malaria in Kahnooj. Then, the accuracy of NDVI (Normalised Difference Vegetation Index) and altitude as predictors of geographical distribution of malaria is assessed (Sections 3.7.5.2 and 3.7.5.3). The last Section (3.7.5.4) concerns the combination of temporal and spatial variations of malaria and their associations with ground-based and remotely sensed.

3.7.5.1. Geographical distribution of malaria

Figures 3-22 to 3-25 illustrate the spatial and temporal distribution of malaria in Kahnooj in the period 1994-2001. The cut of points between low, medium and high incidences were defined in a way to maximise their visual contrasts. Although the risks varied substantially between years and transmission periods, the areas in the centre and south east of district always had the minimum and maximum infection risks respectively. The spatial distribution of malaria in 1998 was more or less similar to the other years and most of malaria cases were reported from five SSDs and the highest risks were observed in the SSD in the south-east of Kahnooj. It suggests that the 1998 outbreak occurred in just a few SSDs.

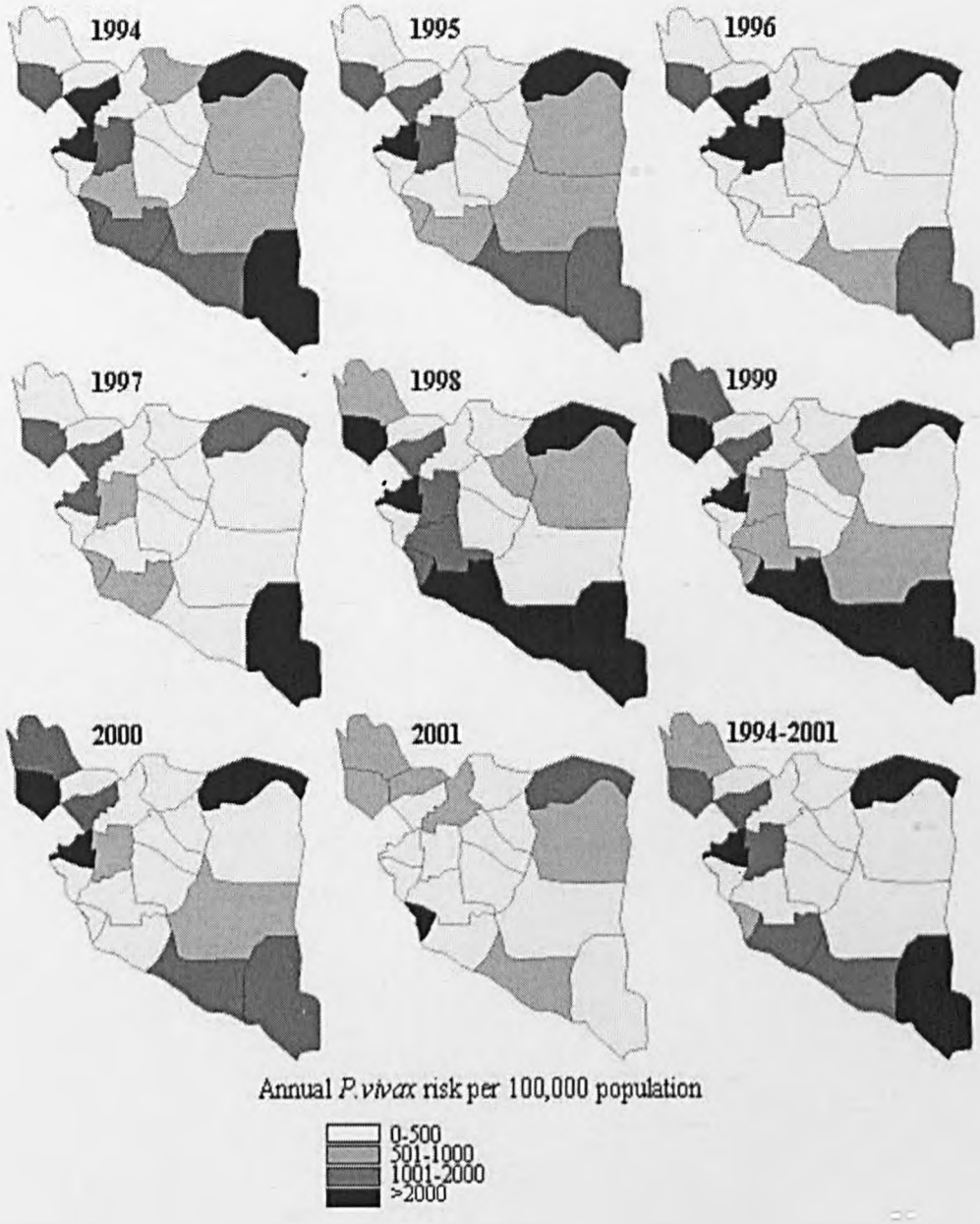

Figure 3-23: The annual spatial distribution of *P. vivax* risk in Kahnooj between 1994 and 2001 by SSD

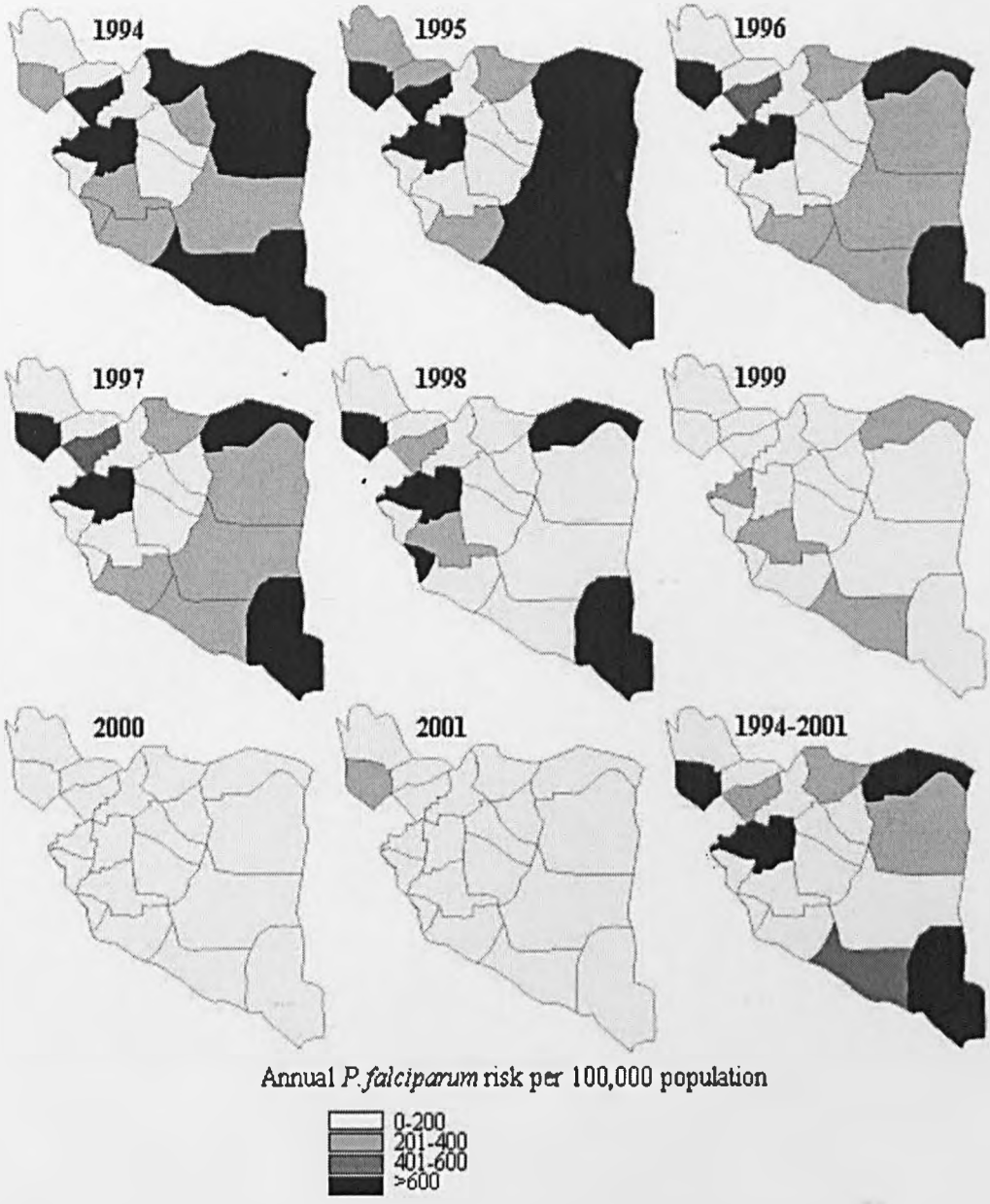

Figure 3-24: The annual spatial distribution of *P. falciparum* risk in Kahnooj between 1994 and 2001 by SSD

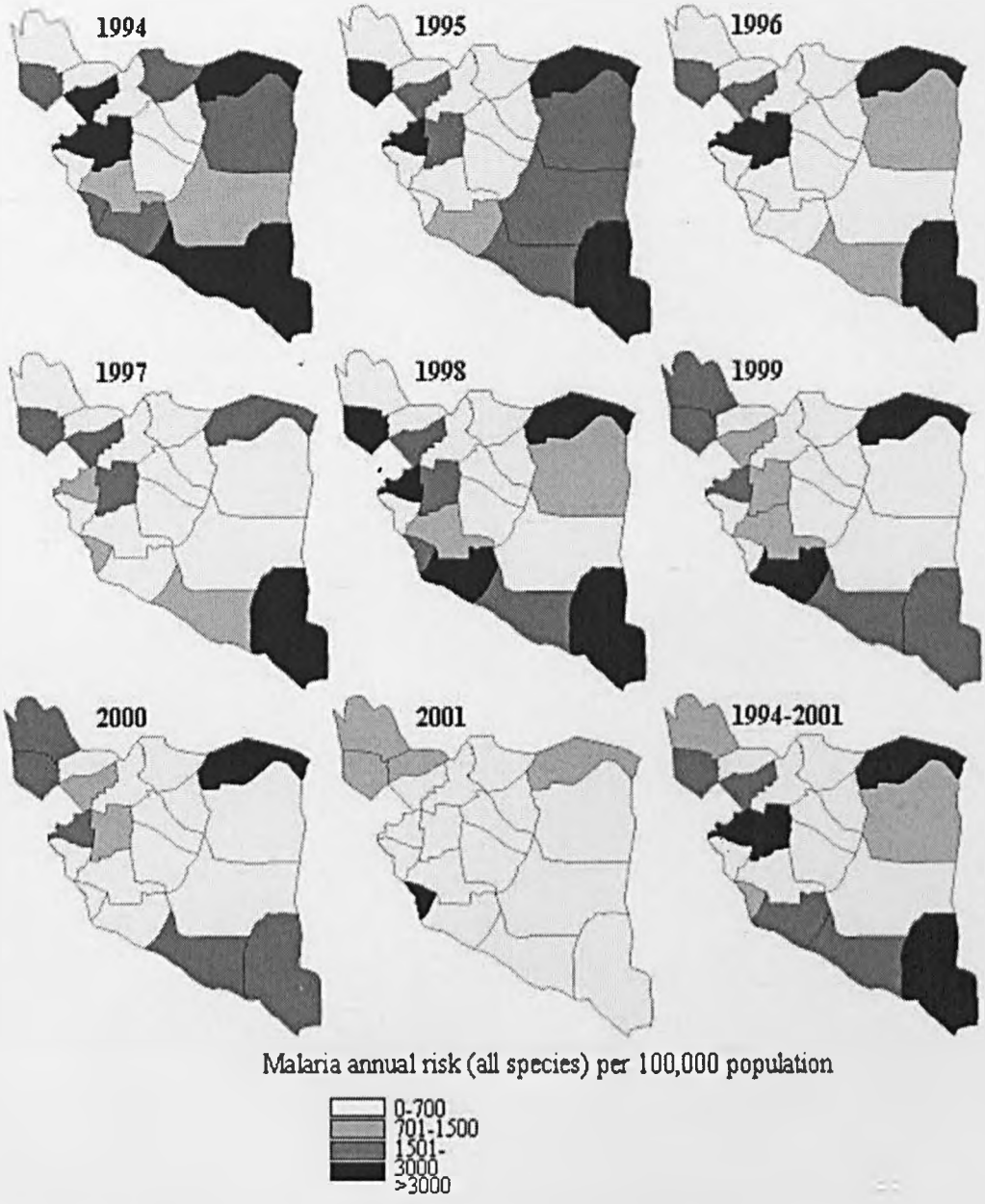

Figure 3-25: The annual spatial distribution of malaria risk (all species) in Kahnooj between 1994 and 2001 by SSD

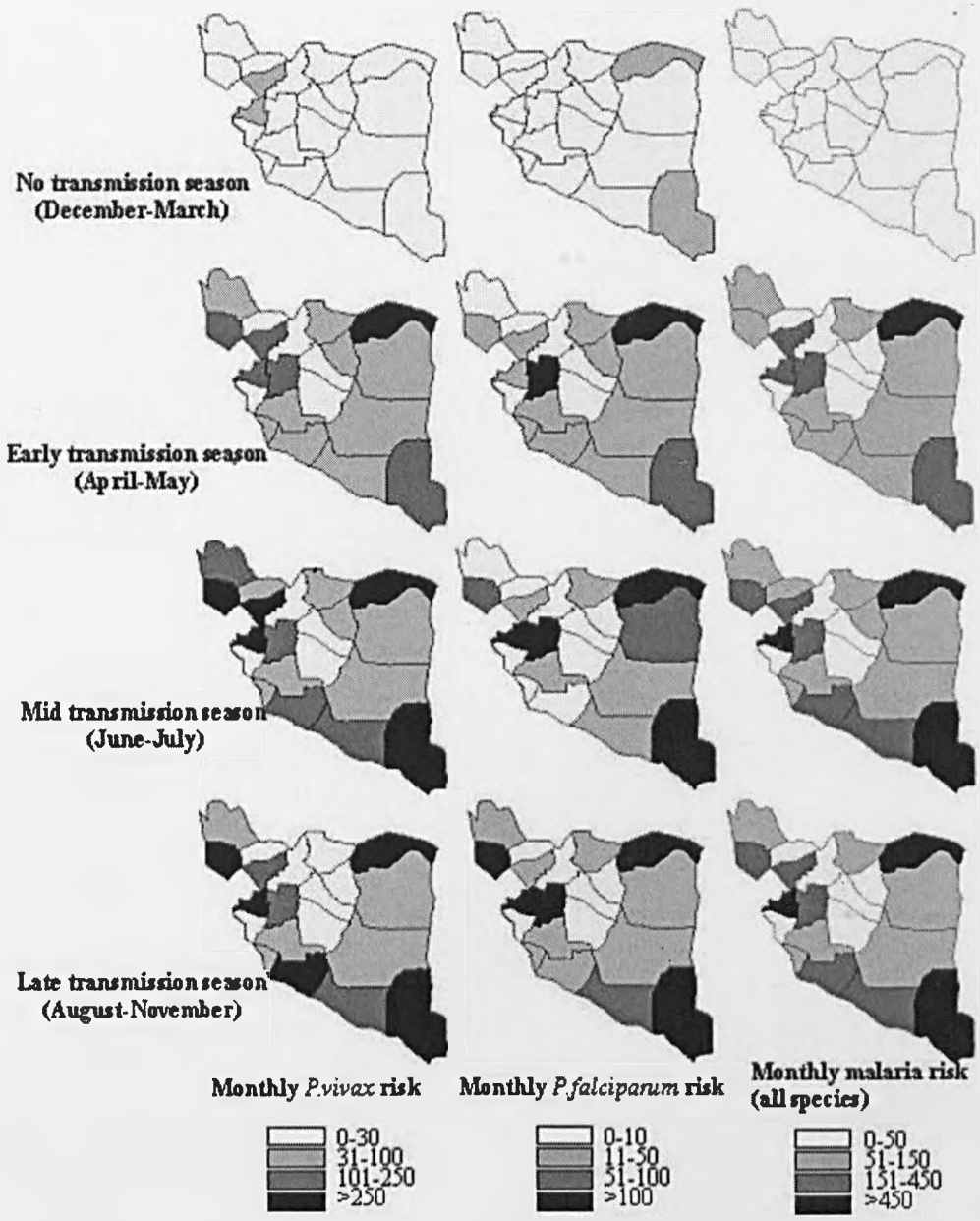

Figure 3-26: The spatial distribution of monthly malaria risk in Kahnooj classified by the transmission period (no transmission, early, mid and late transmission period) between 1994 and 2001 by SSD

3.7.5.2. High spatial resolution NDVI

This section links the risk of malaria cases in villages to their NDVI. Village NDVI were extracted from the Landsat images with 30-meter spatial resolution (Section 3.5.3), which were taken in December 2000 and January 2001. Then, the average of NDVI around each village was computed, from 15m up to 6 kilometre radius. Having done a sensitivity analysis the optimum radius for the mean of NDVI was chosen based on pseudo R^2. Fractional polynomial (FP) Poisson model (m=2) was used to link the number of cases in 2001 to NDVI adjusted for village population (Section 3.4.4).

Modelling NDVI within a five km radius showed maximum pseudo R^2 values (Table 3-10).

Using FP models, the pseudo R^2 values were improved considerably. The pseudo R^2s for the linear effect of NDVI in 5km around each village for *P. falciparum* were 0.09 which increased to 0.15 with the FP model (Table 3-10). The FP models also improved the pseudo R^2 from 0.09 to 0.12 in *P. vivax* and 0.1 to 0.13 in all species. However it is still low compared to the values in previous sections (Table 3-8).

Table 3-10: The pseudo R^2 between malaria risks and the average NDVI around villages in 2001

Radius[1]	All species		*P. falciparum*		*P. vivax*	
	Linear	FR[2]	Linear	FR[2]	Linear	FR[2]
15m	0.004	0.009	0.006	0.07	0.006	0.06
1km	0.04	0.14	0.02	0.04	0.02	0.05
2km	0.07	0.17	0.03	0.03	0.03	0.04
3km	0.08	0.16	0.06	0.07	0.06	0.07
4km	0.08	0.12	0.07	0.1	0.09	0.11
5km	0.09	0.15	0.09	0.12	0.1	0.13
6km	0.06	0.08[3]	0.05	0.09[4]	0.05	0.07[5]

1: The average NDVI around each village was computed in circles with 15m up to 6km radiuses
2: Fractional polynomial, degree two
3: Powers (1,2); 4: powers (-2,-0.5); 5: powers (-2,-0.5))

These models mostly over-estimated the number of cases (Figure 3-26). *P. falciparum* was observed mostly in north-west, north and south-east of Kahnooj. However, the predicted map classified some villages in the west and north-east of the district as moderate to high risk. Furthermore, *P. vivax* was less common in the centre of district than the predicted by the model.

The number of malaria cases was low in 2001. Only 50 *P. falciparum* cases were reported in 2001 compared to more than 2000 in 1994 (Figure 3-14). Although the drop in number of *P. vivax* cases was not as sharp as *P. falciparum*, *P. vivax* was reported in only 8% of villages in 2001. Therefore it was not possible to check model accuracy by subdividing data into modelling and checking parts classified by time and space in 2001.

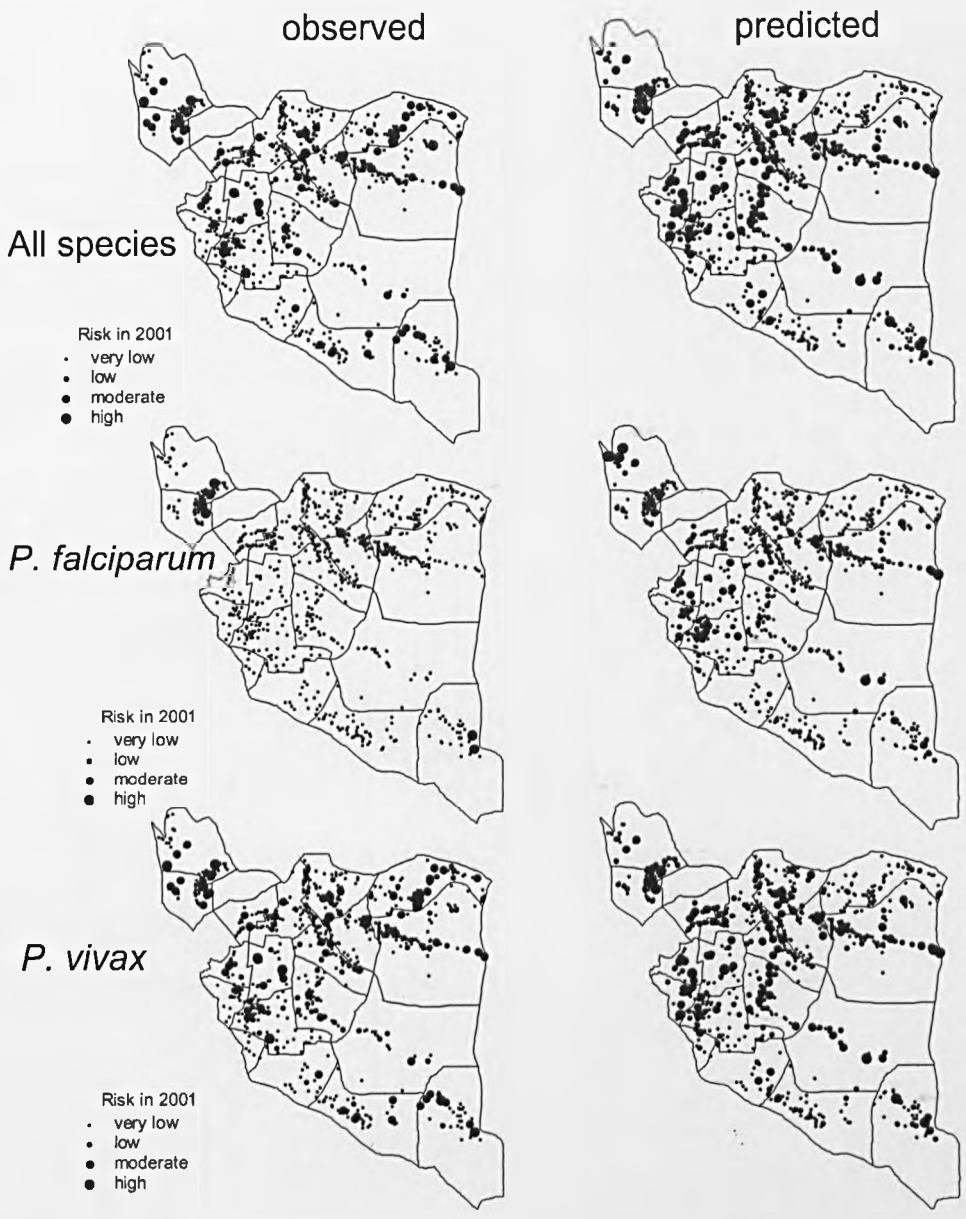

Figure 3-27: The observed and predicted risk maps of malaria in 2001 in Kahnooj, the predicted maps were computed based on NDVI around villages (in 5km radius) (Section 3.7.5.2)

3.7.5.3. Elevation

This section describes the linking of malaria data to altitude. The village altitudes were computed based on the DEM images with 1km resolution.

Then, the mean altitude around each village was computed, from one up to six kilometre radius. Having done a sensitivity analysis the optimum radius for the mean of altitude was chosen based on pseudo R^2. A fractional polynomial Poisson model was used to link the number of cases in 1994-2001 to altitude adjusted for the village populations.

Modelling mean altitude in 2-5km radii showed very close pseudo R^2 values (0.06 in *P. falciparum*, 0.1 in *P. vivax* and 0.09 in all species). Hence, a three kilometre radius was chosen because it is a plausible distance based on the dispersion of living places within villages in Kahnooj and the flight range of mosquitoes.

Using FP models improved the pseudo R^2 values considerably. The pseudo R^2 for the linear effect of altitude in three km around each village for *P. falciparum* was 0.06; the corresponding R^2 in FP model was 0.17 (powers: 0 and 0.5). The FP models also improved the pseudo R^2 from 0.1 to 0.19 (powers: 1 and 1) in *P. vivax* and 0.09 to 0.19 (powers: 0.5 and 0.5) in all species respectively.

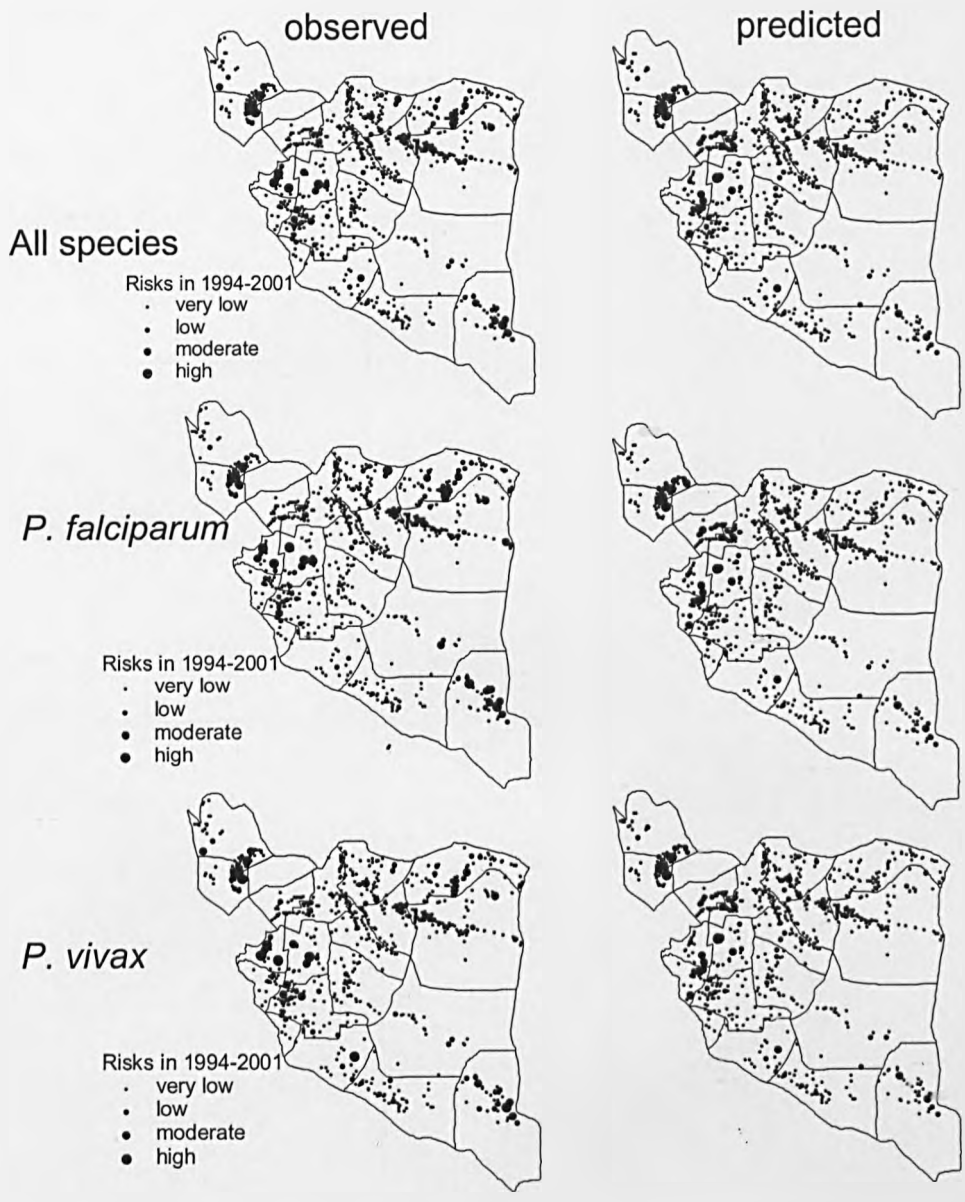

Figure 3-28: The observed and predicted risk maps of malaria in 1994-2001 in Kahnooj, the predicted maps were computed based on the mean of altitude three kilometres around villages by using fractional polynomial models (see text)

The predicted risk maps mostly identified high risk villages particularly in the west part of the district (Figure 3-27). However, they missed most of the high risk villages in the northeast, and some of the high risk villages in the southeast of Kahnooj.

Risk of disease had a non linear association with altitude. Malaria was rare in villages with less than 450 or more than 1400 meter altitude. The maximum risks were observed in villages with 700 to 900 meters altitude. These findings were persistent between years, and species.

3.7.5.4. Temporal NDVI and LST

Land Surface Temperature (LST) is a common remote sensing variable, which is often used as a proxy for ambient temperature. Since vegetation and temperature are significant determinants of mosquito densities, a great deal of research has been done to assess the associations between these indices and malaria *[42,47-49,57]*.

In previous sections, only high spatial resolution NDVI (30m) in December 2000 and January 2001 was used for modelling; however, serial values of NDVI and LST with low resolution ($8km^2$) are used in this section to model spatial and temporal variation of malaria.

This section assesses how much of the temporal and spatial variations of malaria in Kahnooj can be explained by remote sensing data. Using Poisson models, mean monthly NDVI and LST of villages were linked to the malaria risk between January 1994 and September 2001. The remote sensing data were extracted from a set of serial monthly satellite images with 8x8km spatial resolution (Section 3.5.3). The models assess how much NDVI and LST improve the accuracies on top of seasonality, time trend and autocorrelations in risks.

These analyses were carried out at three spatially distinct levels: the village, SSD and district levels. At the village level, the NDVI and LST of villages

were linked to their malaria risks. In the SSD and district levels, the mean of NDVI and LST of villages in each area or in the whole district were linked to the malaria risks.

Pseudo R^2 values were smallest for the village models, higher for the SSD models and highest in the district models (Table 3-11). This is because there is more variation to be explained if the model predicts at the village than the district level. In other words, some of the between village variation was smoothed out of the SSD and district levels.

Having checked the associations from 0-6 month lags a one month lag maximised the pseudo R^2 between infection risks and NDVI and LST in most of the models (Table 3-11). This finding can be explained by mosquito biology and malaria incubation period. Suitable climate conditions are needed for at least one generation to increase the mosquito density and accelerate transmission. Also the incubation period of malaria is around one to two weeks in human body; and most infected subjects contract disease after around ten days. Therefore, one month is a reasonable gap between climate and malaria risk.

In general, *P. falciparum* had stronger associations with NDVI than with LST; on the other hand, the associations between *P. vivax* with NDVI and LST were comparable. It is difficult to explain this finding by the available data, which should be explored in further studies.

Having added NDVI and LST on top of the seasonality and time trend, the pseudo R^2 values increased significantly. The model M10 in Table 3-11 assesses the effects of seasonality, linear time trend and autocorrelation on the risks. Adding NDVI and LST in the model (M11), the pseudo R^2

increased considerably. However, in terms of the numbers and percentages of under and over estimations, the models with remote sensing data did not have better predictions (Table 3-12).

Therefore, it seems that the remote sensing data with such low resolution does not improve the accuracy of models in prediction of malaria (next Section).

Table 3-11: The pseudo R^2 of Poisson models classified by the species based on village, SSD or whole district data

Model number and Explanatory variables		Pseudo R^2								
		P. falciparum			*P. vivax*			All species		
		village	SSD	Distr-ict	village	SSD	Distr-ict	village	SSD	Distr-ict
Models based on remote sensing data										
M1	Current month NDVI LST	0.03 0.03	0.06 0.07	0.07 0.18	0.04 0.04	0.07 0.12	0.13 0.4	0.05 0.05	0.08 0.12	0.12 0.35
M2	Previous month NDVI LST	0.01 0.001	0.06 0.09	0.05 0.22	0.01 0.001	0.04 0.13	0.06 0.44	0.08 0.05	0.06 0.14	0.06 0.39
M3	Two months previous NDVI LST	0.02 0.001	0.03 0.07	0.009 0.16	0.01 0.001	0.008 0.07	0.001 0.24	0.06 0.05	0.02 0.08	0.004 0.24
M4	M2 & M3 NDVI LST	0.02 0.001	0.06 0.09	0.06 0.23	0.02 0.001	0.06 0.13	0.13 0.46	0.09 0.06	0.07 0.14	0.11 0.41
M5	M4 & their quadratic effects NDVI LST	0.02 0.001	0.06 0.1	0.11 0.27	0.02 0.001	0.07 0.14	0.14 0.48	0.09 0.06	0.07 0.15	0.13 0.44
M6	M4 with combine effect of NDVI and LST	0.03	0.13	0.33	0.03	0.16	0.5	0.13	0.17	0.47
Models based on time trend, seasonality and autocorrelation										
M7	Linear effect of year	0.05	0.07	0.07	0.001	0.07	0.13	0.005	0.08	0.12
M8	Sine transformation of time (seasonality)	0.06	0.11	0.26	0.06	0.003	0.44	0.07	0.16	0.41
M9	M7 & M8	0.11	0.21	0.49	0.06	0.16	0.44	0.07	0.18	0.45
M10	M9 & the number of cases in previous month	0.11	0.40	0.65	0.07	0.29	0.66	0.08	0.36	0.67
The final model based on time trend, seasonality, autocorrelation and remote sensing data										
M11	M6 & M10	0.17	0.46	0.77	0.12	0.32	0.73	0.14	0.4	0.75

3.7.6. Summary of the prediction results

The model accuracies were assessed by looking at the number and percentages of over- and under-estimations in the checking part.

The models assessed the accuracies in three spatial distinct levels: whole district, SSD and village levels.

Of all the models, the simplest one extrapolated the number of cases from the previous month's data. In other words, it implied that the number of cases in a month was the simplest predicted value for the number of cases in the following month. As this model is very simple, it does not need any computation. In the field therefore, the health authority is predicting the malaria risk based on this simple prediction. This study aimed to find how much predictors could improve the early warning accuracy. Hence, the results of complex models are comparing with this simple model.

In the next step, the accuracy of models based on seasonality and time trend were assessed. Then, ground climate data were added. In the most complex model, remote sensing variables also were entered. It should be mentioned that in this study, ground climate data for only one point (Kahnooj city) were available. Therefore, the effect of ground climate data was assessed only at the in district level.

Table 3-12 summarises the main results of the models. The village models had the greatest over and under-estimations; which was due to between village variations (Section 3.7.5.4).

In contrast to remote sensing data, ground climate data improved the model accuracies considerably. Comparing the over and under-estimations in the simplest models (extrapolations form the previous month) with other models shows that remote sensing data did not improve predictions. However, ground climate data improved the predictions in district level.

Table 3-12: Over and under-predictions of models based on seasonality, time trend and ground and remote sensing data

Spatial level	Modelling part[1] (%[3])		Checking part[2] (%[3])	
	Over estimation	Under estimation	Over estimation	Under estimation
District data				
Predicted value extrapolated from previous month's data				
P. falciparum	1,020 (23.9)	1,020 (23.9)	372 (27.3)	303 (25.6)
P. vivax	2,118 (24.1)	2,118 (24.1)	438 (22.6)	441 (22.6)
All species	2,994 (23)	2,994 (23)	613 (22.1)	743 (24.7)
Seasonality and time trend				
P. falciparum	1,056 (24.8)	1,056 (24.8)	553 (35.2)	662 (42.1)
P. vivax	1,653 (18.8)	1,653 (18.8)	586 (25.4)	579 (25.1)
All species	2,517 (19.4)	2,517 (19.4)	1,387 (27.5)	1,575 (31.2)
Seasonality, time trend and ground climate data[4] (dekad data, Table 3-9*)*				
P. falciparum	1,113(18.4)	1,113(18.4)	321(16.3)	296(20.1)
P. vivax	1,454(14.4)	1,454(14.4)	408(18.4)	365(17.1)
All species	2,326(14.4)	2,326(14.4)	570(16.5)	581(16.7)
Seasonality, time trend and mean of LST and NDVI[4]				
P. falciparum	796 (18.7)	796 (18.7)	709 (45.1)	376 (23.9)
P. vivax	1,425 (16.2)	1,425 (16.2)	697 (20.0)	812 (23.3)
All species	2,131 (16.4)	2,131 (16.4)	1,271 (25.2)	1,187 (23.5)
SSD data				
Predicted value extrapolated from previous month's data				
P. falciparum	1,796 (40.4)	1,807 (40.6)	535 (38.4)	524 (37.6)
P. vivax	3,171 (34.9	3,580 (39.5	1,286 (40.2)	864 (27)
All species	4,519 (33.6)	4,940 (36.7)	1,654 (36.2)	1,220(26.7)
Seasonality and time trend				
P. falciparum	2,575 (57.9)	2,575 (57.9)	787 (56.5)	767 (55.1)
P. vivax	4,528 (49.9)	4,528 (49.9)	1,163 (36.4)	1,660 (51.9)
All species	6,588 (48.9)	6,588 (48.9)	1,674 (36.6)	2,265 (49.6)
Seasonality, time trend, NDVI and LST				
P. falciparum	2,470 (55.6)	2,470 (55.6)	673 (48.3)	759 (54.5)
P. vivax	4,413 (48.6)	4,413 (48.6)	1,179 (36.9)	1,602 (50.1)
All species	6,424 (47.7)	6,424 (47.7)	1,647 (36.0)	2,215 (48.5)

Continued on next page

Continuation of Table 3-12

village data				
Predicted value extrapolated from previous month's data				
P. falciparum	3,638 (82.9)	3,912 (89.2)	1,233 (84.9)	952 (65.6)
P. vivax	6,480 (69.8)	6,702 (72.2)	2,133 (71.5)	1,903 (63.8)
All species	9,514 (69.9)	10,016 (73.6)	3,137 (70.1)	2,621 (59.2)
Seasonality and time trend				
P. falciparum	3,859 (88.0)	3,859 (88.0)	1,183 (81.4)	1,293 (88.9)
P. vivax	7,232 (77.9)	7,232 (77.9)	2,457 (82.3)	2,342 (78.5)
All species	10,567 (77.6)	10,567 (77.6)	3,441 (77.8)	3,479 (78.6)
Seasonality, time trend, NDVI and LST				
P. falciparum	3,836 (87.4)	3,836 (87.4)	1,205 (82.9)	1,285 (88.4)
P. vivax	7,221 (77.8)	7,221 (77.8)	2,599 (87.1)	2,309 (77.4)
All species	105,34 (77.4)	10,534 (77.4)	3,592 (81.2)	3,424 (77.4)

1: The model was built based on three-quarters of monthly data (modelling data)
2: The fitted value was computed based on the estimated parameters in modelling data
3: Numbers of over or under-estimation divided by total number of cases; since the denominators (the number of cases) were varied between models, the percentages should be used to compare the model accuracies
4: The model is based on the dekad mean of humidity and temperature
Due to the unexplained outbreak of malaria in 1998, the accuracy of all models was dropped considerably. Having excluded the data of 1998, the models predicted the malaria with higher accuracies (Section 3.7.4.6)

In summary, the models predicted the number of cases one month ahead which is a reasonable gap from a practical point of view. The models at the district level had the best accuracies and RS data did not improve the accuracies. Section 3.8.5 discusses the application of these results in the field and presents the possible explanations for lower accuracies in the village models and models with RS data and some suggestions to improve the models.

3.7.7. Local transmission

This section explores the risk factors of local malaria transmission. Local transmission risk map can be a useful tool for health system to identify and

concentrate its activities in high risk areas to prevent outbreaks. In other words, in terms of early warning system (one month ahead), local transmission risk map may be as important as the risk map based on incidence risks.

The local transmission in each village was defined as the presence of at least two conspecific malaria cases in a month or one conspecific case in two consecutive months. The effects of NDVI and LST were assessed by comparing the areas under the Receiver Operating Characteristic (ROC) curves. Furthermore, having adjusted for seasonality and within village clustering, the effect of LST, NDVI, population and history of the disease in the village were assessed.

ROC is a graphical representation of the trade off between the false negative and false positive risks. It is the standard approach to evaluate the sensitivity and specificity of diagnostic procedures. Each point on the ROC curve is associated with a specific diagnostic criterion. The area under the ROC curve has become a particularly important metric for evaluating diagnostic procedures because it is the average sensitivity over all possible specificities. The more discriminatory curves are those which go further towards the top left corner (Figure 3-28).

Villages were classified as either positive or negative in each month for each species. They were considered positive if at least two species-specific cases were reported in a village in a month or in two consecutive months; i.e., two cases in one month or one per month for two consecutive months. Also, due

to the high number of imported cases in cities, data from cities were not included in this analysis.

The data were classified into modelling (75%) and checking parts (25%) randomly. To check the accuracy of predictions, models were run in the modelling data. Then, local transmission risks in the checking data were computed given the fitted equations. The model accuracies were estimated based on the differences between observed and predicted values in the checking part of the data.

Logistic regression was used to model the local transmissions for *P. falciparum*, *P. vivax* and any species. Having taken into account within village clustering, the local transmissions were modelled based on seasonality, time trend, population and history of disease with the same species in the village between 8 and 18 months ago (the highest risk of *P. vivax* relapse was observed in with this gap, Section 3.7.3.2), with or without NDVI, LST. Then, the effects of NDVI and LST were checked by comparing the area under the ROC curves.

All of the explanatory variables showed significant p-values. The odds ratios for the history of diseases were 5.3 and 3.5 in *P. falciparum* and *P. vivax* respectively. Also, there was a significant linear trend between population and the risk of local transmission; the odds ratio increased 1.15 times for every 100 people increase in population (the maximum number of village population was 980 people).

The result of modelling and checking parts are similar (Table 3-13). The largest differences between the accuracy of models with and without NDVI and LST were observed in the *P. falciparum* models in modelling data

(specificity: 75.6% versus 80.1%, area under the ROC: 0.85 versus 0.87 in modelling and checking parts respectively). The other differences were even smaller than this difference. Small differences between the results of checking and modelling parts may imply that the models were predicted the local transmission risks in checking part as accurate as in modelling part.

Including of NDVI and LST increased the model accuracies significantly ($p<0.02$), as well as the area under the ROC curve ($p<0.001$) (Table 3-13). The magnitudes of differences were not large; however, in terms of actual numbers even these small differences, drops the numbers of false positives and negative villages considerably.

Spatial distributions of local transmission risks were more accurate in the *P. vivax* model than *P. falciparum* model (Figure 3-29). Comparing the corresponding maps shows that all models over-estimated the frequencies of local transmissions. However, *P. vivax* model discriminated particularly those villages with low and intermediate risks more accurately than the *P. falciparum* model.

Table 3-13: The accuracy of local transmission models with seasonality, time trend, population and history of the disease, with and without NDVI and LST

	Modelling part			Checking part			All data		
	Pf	*Pv*	All	*Pf*	*Pv*	All	*Pf*	*Pv*	All
Models without NDVI and LST									
Sensitivity	77.1	76.8	72.7	78.2	76.5	72.2	75.9	70.7	71.4
Specificity	75.6	71.5	76.7	80.1	71.4	77.3	75.7	81.8	76.8
Area under the ROC	0.85	0.837	0.838	0.87	0.838	0.833	0.833	0.863	0.832
se of area under the ROC	0.006	0.004	0.004	0.01	0.008	0.007	0.003	0.002	0.001
Models with NDVI and LST									
Sensitivity	78.6	77.1	78.1	78.7	76.8	78.2	78.4	78.3	77.5
Specificity	80.1	74.2	82.2	82.5	74.7	82.1	75.6	81.8	77
Area under the ROC	0.87	0.839	0.864	0.857	0.839	0.86	0.846	0.863	0.845
se of area under the ROC	0.006	0.003	0.003	0.009	0.005	0.005	0.003	0.001	0.001

Pf: P. falciparum, Pv: P. vivax

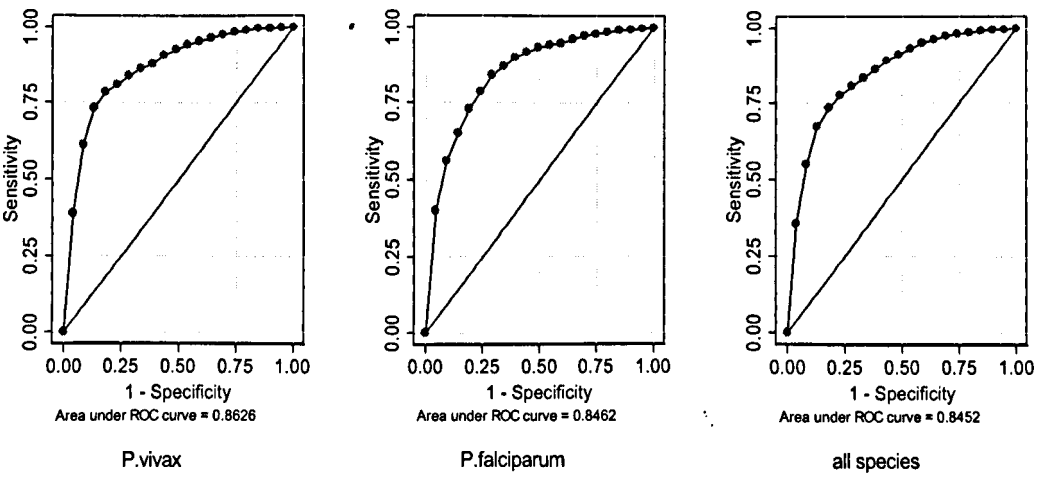

Figure 3-29: Species-specific ROCs, they assess the relationship between sensitivity and specificity of the full models (with NDVI and LST) in predicting local transmissions in all data

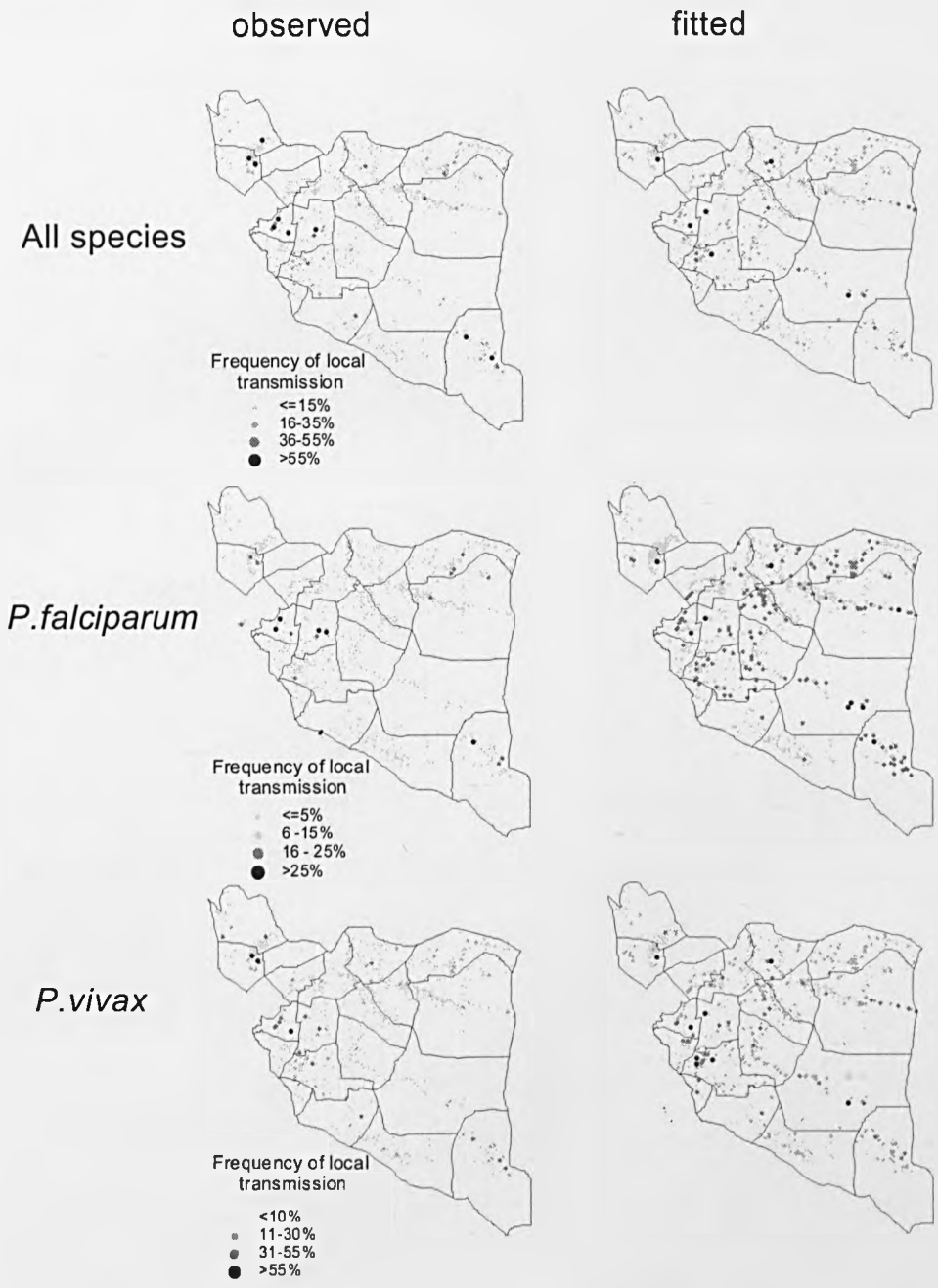

Figure 3-30: Comparing the fitted and observed risk maps of local transmission, the fitted values were computed based on seasonality, time trend, history of disease, NDVI and LST (see text for further explanations)

(Legends are showing the percentages of months were each village had local transmission)

3.8. Discussion

This chapter presents and analyses the malaria data of Kahnooj district, with the aim of assessing the feasibility of an early warning system based on meteorological and remote sensing data in predicting of malaria. The transmission was seasonal and the peak of malaria observed in middle to late summer.

Sections 3.2.3.2 and 3.2.4 explained the situation of malaria in Kahnooj and described the malaria surveillance system. Also, Section 3.7.1 showed that the microscopy results had enough accuracy to be used in these analyses.

The sample size of this study was quite enough to run models: 18,268 malaria attacks were recorded between 1994 and 2002. Although there were relatively few cases, particularly for *P. falciparum* in some time bands, overall, the sample size was enough to model.

Table 3-5 showed the main risk factors of malaria in Kahnooj. Although, Afghani immigrants are mostly blamed for malaria epidemics, this was not supported by the findings. Section 3.7.2 discussed possible explanations for this finding; in particular, that Afghani immigrants are mostly asymptomatic carriers. Therefore, from a public health point of view, it seems that they should be examined more closely and get full treatment, even those who do not contract clinical malaria.

It should be mentioned that to have an outbreak, suitable environmental factors are not sufficient. In central and western of Iran, malaria transmission is very low, but in some villages vectors have effective contact with people (such as Fars province in west of Kahnooj). That is because at

least a few carriers are needed to trigger the outbreak under appropriate conditions. Therefore, screening the carriers among Afghani people may play an important role in controlling malaria in Kahnooj.

Section 3.7.3 discussed the pattern of repeated episodes among cases. It seems that treatment failure, particularly in *P. falciparum* is a very important issue to which policy makers need to pay more attention in order to monitor this rate and mount a timely reaction to its level (Section 3.7.3.1).

This study estimated the relapse rate of *P. vivax* based on a new approach which allows differentiating the rate of real relapse from re-infection (Section 3.7.3.2). The observed relapse pattern of *P. vivax* in Kahnooj was comparable with the pattern in other temperate areas *[104]*. Also, comparing the relapse rate with other studies *[106]*, it seems that anti-relapse treatment had an acceptable effectiveness in Kahnooj.

3.8.1. Temporal variations

Except the outbreak in 1998, malaria had a decreasing trend between 1994 and 2002 in Kahnooj, which can not be explained by meteorological factors (Figure 3-15). Based on these analyses, there is no simple explanation for the 1998 outbreak, but based on the national statistics, it is clear that the outbreak occurred in some areas in south-east of Iran. Some experts in national and provincial levels explained it by the unstable situations inside Afghanistan which increased the number of illegal immigrations to Iran in 1997 and 1998. Afghanistan's ruling Taliban militia massacred thousands of civilians, including nine Iranian diplomats and about 3,000 Hazaras - a Persian speaking Shi'a minority - when it seized the northern city of Mazar-i

Sharif in August 1997. This caused a significant influx of refugees mostly to Iran in the end of 1997 and early months of 1998, because most of Shi'a Afghanis preferred to immigrate to Iran. There are no accurate data concerning the number of immigrants.

The decreasing level of malaria might be due to intensive control programmes in the field. In 1979, Iran had a deep political revolution; after that Iran had a long war with Iraq for 8 years, which absorbed a considerable part of Iranian national resources. Since 1990, Iran has had a more stable political and economic situation, which allowed the government to allocate more resources for health and education. Now more than 92% of inhabited villages in the Kahnooj have access to electricity, while in 1979 just Kahnooj city had electricity. The illiteracy rate has dropped sharply in the last two decades, and now is less than 15% among people with aged 10 years or more. Therefore, the decreasing trend of malaria could be due to the improvement in socio-economic situations which it is expected to continue.

Furthermore, severe drought in the last decade might have a considerable impact on the epidemiology of malaria in Kahnooj not only directly but also via changing the life styles of people.

Malaria had very clear seasonal pattern in Kahnooj. Seasonality alone explained around one third of the temporal malaria variations; also the combined effect of time trend and seasonality explained around sixty percent of the variation (Table 3-8). Seasonality alone explained more variation for *P. vivax,* while the combined effects of seasonality and time trend explained more variation in *P. falciparum.*

3.8.2. Spatial variations

Malaria risk had a wide variation in Kahnooj district (Figure 3-25). Vast areas of Kahnooj with considerable variation in climate can be counted as the main source of the spatial variation in malaria risk. In addition, malaria was more prevalent in very remote villages, mostly in south-east of the district, which have more illegal Afghani immigrants and lower socio-economic situations.

There was an interaction between spatial and temporal distributions of malaria. The seasonal pattern of malaria was not exactly the same in subsubdistricts (Figure 3-24). Also, the annual risk of malaria had different patterns; for instance, although the SSD in south-east of the district had the highest risk in most of the years, in 1999-2001 its risk was considerably lower than the other high risk SSDs.

Nonetheless, the middle part of the district had the lowest risks of malaria in most of the years and seasons.

3.8.3. Species-specific variations

There were considerable differences between temporal and spatial variations of *Plasmodium* spp. Section 3.7.4.1 explained the differences between *P. vivax* and *P. falciparum* annual curves; the former species had a bimodal curve with concavity in the mid summer, while the latter usually showed a sharp peak at the end of transmission season.

This dissimilarity could be due to the different sensitivity of species to meteorological factors, particularly temperature. Figure 2-1 showed that *P. falciparum* requires higher temperature to mature within the mosquito (sporogonic cycle) as fast as *P. vivax*. Therefore, in the beginning of the hot season, *P. vivax* peaks faster than *P. falciparum*.

Unsuitable meteorological conditions are the most plausible explanation for concavity in the *P. vivax* curve in mid summer which has been reported in other studies *[37,123,124]*, *P. vivax* transmission started earlier in summer. However, in mid summer the temperature and humidity have their maximum and minimum values respectively. These conditions might decrease the survival rate of mosquitoes and slow down the transmission rate, and thereby reduce caused a drop in the *P. vivax* incidence.

The differing sensitivities of *Plasmodium* spp to meteorological factors, and the wide variation in these factors in Kahnooj, means that the spatial variations of species were not consistent. Comparing Figure 3-22 and Figure 3-23 illustrates these differences and suggests the need for species-specific models.

3.8.4. Accuracy of final models

Final model results are summarised in Table 3-12. This table showed the number and percentages of over and under-estimations of predictions one month ahead. Predictions relate to three spatially distinct levels: district, SSD and village levels.

The outbreak of malaria in 1998 decreased the model accuracies considerably. For example in the district level models, ground climate data

improved the predictions around 6-8%. Having excluded the data of 1998, these differences rose to 18-23%. However, it should be mentioned that all of the analyses and conclusions were based on the whole data including the 1998 data. Therefore, ground climate data can improve the accuracy of early warning models.

On the other hand, remote sensing data did not improve the model accuracies considerably, which could be due to the following reasons:

1. The resolution of satellite images was 8x8km; the extracted data from these coarse images might not have enough accuracy to demonstrate any associations.
2. Kahnooj is arid and most of its area is desert; therefore, NDVI (as the one of most common vegetation indices) might have not enough accuracy to measure the real greenness of land in arid and semiarid areas comparing to the other vegetation indices.
3. The cold cloud duration (CCD) was not available in this study; therefore, rainfall was not entered in these models. Rainfall had the strongest association with the annual risk of malaria (Section 3.7.4.1), but was not estimated in the remote sensing data. Therefore, one of the most important explanatory variables was absent.
4. The remote sensing data reported monthly average of NDVI and LST; monthly averages might not have enough accuracy to present the situation in the whole month.

3.8.5. Practical application of the models

These models show that ground-based meteorological data can be used as a very simple and accessible tool to predict malaria risk. As mentioned in the previous section, ground-based data can improve the accuracy of predictions by around 10%. Nonetheless, the predictions should be species-specific.

Using GIS to create risk maps may help the health system to identify the SSDs at highest risk to mount more intensive control activities. A director of malaria programme in the district level can predict the high risk SSDs one month ahead and concentrate resources to prevent outbreaks. Since risk varies in time and space, a mobile control team may be most efficient.

The final predicted model, based on ground climate data, improved the accuracies around 10%. Predictive variables for these models are readily available in the field, so an improvement of even a few percent makes them feasible.

3.8.6. Local transmission

Identifying the risk of local transmission is very important in control and eradication programmes. Section 3.7.7 described the risk factors of the local transmissions and the role of NDVI and LST.

The sensitivities and specificities of the models were high and their areas under the curves were highly significant. NDVI and LST increased the area under the curve around one percent. Even without NDVI and LST, the history of malaria and population were very important predictors. Therefore,

these models can provide useful information about the local transmission risks. These risk maps together with the risk maps of malaria intensity can help the health system to make more accurate decision in identification of the high risk villages.

3.8.7. Limitations and suggestions for further studies

As for any other surveillance data, the accuracy of data was a point of concern in this study. Although, there was some convincing evidence about the quality of data, there might still be some errors or mistakes in case finding and reporting.

The models could not explain the outbreak in 1998. It seems that some other important factors played important role which they were not measured in this study. Further studies are needed to access the impact of other factors and address specifically to the outbreak causes.

These results are based on eight year surveillance data; during these years severe drought might change the pattern of malaria transmissions and life styles of people. A longer study might explain time trend of malaria variation more accurately. For this type of study, even monthly malaria data within subdistricts may be enough.

The results of this study illustrate the application of GIS in malaria control programmes. To introduce GIS in practice, health policy makers should be aware of its usefulness; therefore, researchers should simplify their findings and write their reports as clearly as possible. Otherwise, the heath system will not accept that this new tool could modify its traditional data collecting method and reporting system.

The results generally do not show any potential for the use of remote sensing data in predicting malaria in Kahnooj. Section 3.8.4 discussed some explanations. Therefore, further studies are needed to address clearly on the application of remote sensing data. These studies should assess the association between malaria risk and remote sensing data in finer scale and include an estimation of rainfall.

The ground climate data were from only one point (Kahnooj city) which limited the climate models to only districts level. The results showed that ground climate data were predicted the number of malaria cases more accurately than remote sensing data. Therefore, it may suggest that more attention to the data climate stations could improve the feasibility of early warning system.

Another limitation of this study was the definition of local transmission which was the simplest possible definition. Nonetheless, another study with more accurate definition of local transmission may find stronger evidence about the feasibility of such a model in prediction of malaria epidemics.

3.9. Final conclusion

The main objective of this study was to assess the feasibility of an early warning system based on ground and remote sensing data. Based on the accuracies of the models, it seems that models based on low spatial resolution remote sensing data (NDVI and LST) are not feasible. However, more studies are needed to assess the effect of remote sensing data in finer resolution with more predictors such as rainfall; and quantify the feasibility.

Since these types of data are expensive, a cost-effectiveness analysis may be needed.

Ground climate data (which are available free of charge) improved the model accuracies and it seems that early warning system based on these models is feasible (Section 3.8.4).

Furthermore, combination of risk maps of malaria incidence and local transmission may improve the feasibility of early warning system. However, it is suggested that other studies explore the feasibility of models based on multiple climate centre data.

Part Two: Interaction between *Plasmodium* spp

Prospectus

This part is about the interaction between *Plasmodium* species and has four chapters. The first one (Chapter 4) reviews the literatures and discusses the possible sources of the interaction.

The next chapter (Chapter 5) is a systematic review of the literature and meta-analysis to assess the interaction between *P. falciparum* and *P. vivax* and explore the potential source of heterogeneities. Many papers have reported the frequency of mixed infections, and there are a few review papers *[125-128]*. However, neither a systematic review nor a meta-analysis has been published about the interaction between *P. falciparum* and *P. vivax* and the source of heterogeneities.

This meta-analysis explores the effects of age group, the presence of fever, frequencies of individual infections, the geographical location of studies, and the temporal and spatial spans of the studies on the interaction between *Plasmodium* spp. Age group and presence of fever are related mostly to possible biological and immunological pathways while the temporal and spatial spans and geographical locations associate mainly to the environmental factors.

Chapter 6 presents the results of a mathematical model; it evaluates the impact of heterogeneity in infection risk on the interaction between species.

The last chapter (Chapter 7) explores the interactions between *Plasmodium* spp in the Garki data *[129]*, one of the largest epidemiological studies on

malaria, and examines the effect of study type in order to explain the differences between cross-sectional and longitudinal findings.

In contrast to cross sectional studies, the number of longitudinal studies among publications is low. In addition, no large epidemiological data set has been analysed in both ways to assess directly the effect of study type.

CHAPTER 4

4. Introduction

Dual infections with *Plasmodium* spp are common and reported in many studies. Results of these studies generally do not support the hypothesis of species sampling independence *[130]*, and a wide range of associations has been reported. In addition, there is an inconsistency between the findings of cross-sectional and longitudinal studies.

Exploring the interaction between species may help us to understand more about the biology of *Plasmodium* spp within the human body and to extend current knowledge about the immunological mechanisms against malaria.

4.1. Definition of interaction between species

In this study, the interaction between *Plasmodium* spp is assessed by looking at the presence or absence of infections in blood slides. The study utilises measures based upon the difference between the observed number of mixed infections in blood slides and the expected number if infection with one species is independent of infection with other species. The magnitude of interaction is shown by odds ratios.

Based upon the above definition, a positive interaction would imply that mixed infection is more common than expected, and is illustrated by an odds ratio greater than one.

On the other hand, a negative interaction would imply that mixed infection between species is less common that expected; i.e., an odds ratio less than one.

4.2. Background

Human *Plasmodium* spp share the same transmission route. In addition, the exposure risk to mosquitoes is usually positively skewed within populations; i.e., a subgroup of people is more highly exposed to mosquitoes compared to the average population. Therefore, based upon the above two facts, a positive correlation between these infections might be expected *[131]*.

Pinto et al. (2000) found a positive association between *P. falciparum* and other *Plasmodium* spp. They explained their findings by heterogeneity in exposure to mosquitoes. All of their febrile cases had single *P. falciparum* infection; therefore they suggested that that mixed infections may protect against clinical symptoms of disease *[132]*.

Nonetheless, according to the results of most prevalence surveys, fewer mixed-species infections were observed than would be expected based on the product of the frequencies of individual species. This finding suggests that one parasite has excluded another or suppressed its parasitaemia to undetectable levels *[6,133]*.

McKenzie and bossert (1997) reviewed the point prevalence of *P. falciparum* and other *Plasmodium* spp in 35 papers published between 1984 and 1995 *[126]*. They showed that high overall prevalence of infection was associated with significant deficits of dual *P. falciparum* and *P. vivax* infections, while this was not observed in dual *P. falciparum* and *P. malariae*. Since the prevalence described frequencies of associations, they concluded that a pattern of more specific biological interaction, for instance, at the levels of immunity, pathogenicity, or transmission was involved but they did not explain any specific pathways.

Only a few longitudinal studies exploring this issue have been conducted. One conducted by Bruce et al (2000) *[134]* reported that the frequency of mixed infection was very close to the expected number, assuming no interaction.

A study in Thailand showed that *P. vivax* developed in a third of patients treated for acute *P. falciparum* within a month of receiving a regimen containing quinine or quinidine or within two months of receiving mefloquine treatment. However, less than one percent of patients in Thailand presenting with acute malaria were reported to have a mixed infection. The author concluded that either the routine microscopy could not detect most of the mixed infections or acute *P. falciparum* infection had suppressed the blood stage of *P. vivax [135]*.

The effects of interaction terms vary in different studies, and cannot be explained solely by cross immunity. Howard et al. (2001) *[125]* used a log linear regression model to show the association between multiple species parasite infections. They showed that the logarithm of the odds ratio

between *P. falciparum* and *P. vivax* varied over a wide range from –5.08 (in Bangladesh) to +2.56 (in Sierra Leone). In addition, they found that for Asian countries the associations were largely negative; however, positive associations were seen in Tanzania, Papua New Guinea and the USA.

Experimental studies in animals and some epidemiological studies have provided evidence that infection dynamics (the blood stage of parasites and their gametogenesis rates) are affected by cross-species immunity. The majority of papers explained these interactions by cross immunity between species, density-dependent regulation, and differential growth and clearance rates of individual parasite populations resulting from clonal antigenic variation. Most of them focused on the suppressive effect of *P. falciparum* on the dynamics of other species *[135,136]*.

An epidemiological study of the morbidity of malaria in young children in a highly endemic area of Africa suggests that clinical immunity depends mainly on the extent of exposure to blood-stage antigens *[94]*. Mason et al. (1999) *[137]* built a mathematical model to describe only the blood stage dynamics of mixed infection and the effect of immune response to these interactions. However, due to uncertainty about the specific and non-specific-immunity proliferation and capture/removal rates, the model output was examined over a very wide range of values. The findings of this study were inconclusive.

4.3. Possible explanation for positive interaction

4.3.1. Similarity in transmission routes

Similarity in transmission routes is the simplest explanation for the positive associations between the infections. Since all *Plasmodium* spp are transmitted by the same vector, any differences in exposure between persons can automatically be applied to all species *[131]*. In other words, exposure to mosquitoes might increase the risk of transmission of all species simultaneously.

4.3.2. Higher susceptibility of a subgroup of people

It is also theorised that certain subgroups have a higher susceptibility to infections. Based on this theory, susceptible people might get all infections more frequently than the others in the general population. The susceptibility might be due to genetic factors, or lower acquired immunity *[138]* or other factors such as the immunosuppressive effects of chronic infections, coexisting diseases or malnutrition *[139,140]*.

4.4. Possible explanation for negative interaction

4.4.1. Suppression

The suppression hypothesis is supported by data derived from the simultaneous inoculation of two *Plasmodium* spp into laboratory animals. Many studies have shown that one or both species are suppressed. This

may be mediated by competition for host cells or nutrients, or by heterologous immunity. However, the suppressed species rebounds after the other species has abated, and may show a prolonged infection *[135,136,140]*. Fox and Strickland (1989) discarded the suppression effect of *P. falciparum* on *P. vivax* in their studies in Punjab *[141]*.

4.4.2. Cross immunity

It seems that cross immunity does not have any substantial effect on the interaction between species; however, its role cannot be ignored completely. Most of the experimental studies in animals showed that cross-immunity between species did not give considerable protection *[90,142]*. In addition, Bruce et al. (2000) showed that *Plasmodium* spp are independent in humans *[134]*. The effective acquired immunity against each species was low over the long-term; therefore, there is not any convincing evidence to support substantial protective cross-immunity between *Plasmodium* spp *[138,140]*.

4.4.3. Differences in the biology of *Plasmodium* spp

Plasmodium spp have different biological characteristics. For instance, maturation of the pre-erythrocytic hepatic stage of *P. falciparum* is more rapid than that of *P. vivax*, and asexual *P. falciparum* parasites will therefore appear earlier in the peripheral blood after simultaneous inoculation. Thus if patients present very early in the course of infection, low *P. vivax* parasitaemia could easily be missed *[135]*. In the same way, biological differences between *Plasmodium* spp such as relapse, incubation

and infectious periods may explain some part of the observed interaction between species in cross sectional studies.

4.4.4. Environmental factors

Calculation of the expected number of dual infections might be over-estimated based on the product of species-specific prevalence (incidences), particularly in large studies with non-homogeneous temporal and spatial distributions. Therefore, seasonal and spatial variation of these two species might partly explain the apparent interaction between the *Plasmodium* spp *[138,141].*

4.4.5. Missed mixed infections in blood slides

In the assessment of the frequency of dually infected slides by microscope, observational bias in reading films is usually a very important issue. The effect of this bias is more prominent if junior microscopists without supervision read blood films. In addition, the reading method of the films is also crucial. The flexible method has the highest bias. In this method, the reader examines flexible number of fields based upon his/her own judgment. The assessment is usually terminated by finding some positive fields.

CHAPTER 5

5. Systematic review and Meta-analysis

This systematic review of the literature and meta-analysis estimates the interaction between *P. falciparum* and *P. vivax* and explores the possible sources of heterogeneity such as the age group of subjects, geographical and temporal spans of studies and the prevalence of malaria in the populations.

5.1. Objectives

1. To quantify the interaction between *P. falciparum* and *P. vivax*
2. To decide whether or not there is heterogeneity in the interaction terms among the results of papers.
3. To assess the source of the heterogeneities, by focusing on the effects of endemicity, temporal and spatial spans of studies and the age group of the samples

5.2. Data collection method

The main electronic databases in medicine and public health were searched with wide key words to optimise the sensitivity of data collection. Then the

abstracts of selected papers were reviewed. Based on the abstracted data and a proforma checklist, the data were categorised into three groups: not eligible for meta-analysis, eligible for meta-analysis and a group in which their full texts were needed for final decision on their eligibility (Section 5.2.4). In the next step, the full texts of the second and third groups were reviewed and the required data for meta-analysis were abstracted to a standardised form (Section 5.2.5).

5.2.1. Database search method

Medline, Embase and CAB-Health were searched. Due to differences in the formats of these databases, the details of the searching method are explained separately.

5.2.1.1. Medline

This database was searched from the first of January of 1966 to the end of May 2001.

Using the thesaurus of "*Malaria*", the subheading of "*epidemiology*" was selected. These key words selected 3078 papers.

Then this search was limited by adding "*TG=HUMAN*". In this stage, 2977 papers were selected ("*TG*" stands for Target Group).

Then the "*falciparum*" and "*vivax*" words were added. The number of matched citations was reduced to 395.

Accordingly, the final searching phrase was:

"(malaria/ epidemiology) and falciparum and vivax and TG=HUMAN"

5.2.1.2. Embase

This database was searched via the BIDS web site from the first of January 1980 to the end of May 2001.

In the first stage, different expressions were searched separately, and then their results were merged. The search phrases and their results were as follow:

"Malaria/ focus or malaria as key word" 17242 citations

"Epidemiology/ expand" 322657 citations

"falciparum/ malaria or plasmodium/ focus" 7160 citations

"vivax/ plasmodium/ focus" 733 citations

The number of papers with "&" combination of the above phrases was 77.

Result of the final search was saved in "Reprint/ medlars" format.

5.2.1.3. CAB-Health

This database was also searched via BIDS web site from 1st of January of 1973 to the end of May 2001.

The search was started with the keyword "malaria" as the simplest expression.

Then, by adding new key words, the search was restricted. The details of search were as follow:

"*Malaria*" 23,895 citations

"*Malaria & falciparum*" 10,816 citations

"*Malaria & falciparum & vivax*" 1954 citations

"*Malaria & falciparum & vivax & epidemiology*" 577 citations

"(*Malaria/ in key words) & falciparum & vivax & (epidemiology/ in keywords*)" 455 citations

5.2.2. Merging the result of searches

Using Endnote software, the results of the searches were merged. The results of the Medline and Embase searches were imported to Endnote directly. After merging these two files, 11 duplicate papers were found. After deleting the repeated citations, the merged file had 461 papers.

Endnote version 4 had no importing filter compatible with the structure of CAB-Health text file. After generating an appropriate filter, the results were imported to Endnote.

Then, the duplicate papers in the "Medline + Embase" file and CAB-Health file were detected according to their titles, authors and source of paper fields. In this stage, 46 papers were found. After deleting them, only the title was used to mark duplicate papers. The abstracts of all detected citations were checked one by one to minimise errors. In this step, 30 citations were deleted. At the end, 829 citations remained in the main database.

5.2.3. Exporting the citations to Access

The content of the Endnote file was exported in tab-delimited format. Then a new database was created in MS-Access 97 and the content of the file was imported as a table into MS-Access.

An interface was created to show the main bibliography of each paper (authors, publication year and abstract) and some new fields in relation to the content of abstracts in each page. This interface facilitated reviewing abstracts and checking their eligibilities (see next Section).

5.2.4. Reviewing the abstracts

Using a checklist, information about study objectives, sampling methods, type of study and main findings were extracted from the abstracts (appendix one). Based on this information some ineligible papers were excluded. The eligible studies were those ones which reported the frequencies of *P. falciparum*, *P. vivax* and mixed infections among random samples in a define population.

Out of 829 abstracts, 104 (12.5%) papers were categorised as suitable for meta-analysis; also 68 (8.2%) papers were recruited to check their eligibilities by reviewing their full texts. Those papers which reported the results of surveillance systems, chose non random samples such as including immigrants from endemic areas, or estimated neither the incidence nor the prevalence of malaria were excluded (Table 5-1).

Table 5-1: The result of reviewing the abstracts

	Frequency	*Percentage*
Objectives		
Interaction between species	18	2.2
Epidemiology	431	52.0
Control programmes	23	2.8
Entomology	64	7.7
Treatment and drug resistance	37	4.5
New techniques	30	3.5
Others	226	27.3
Sampling method		
Random (any sampling method)	177	21.3
Non-random	125	15.1
Not mentioned	53	6.4
Irrelevant[1]	474	57.2
Did they include immigrant patients?		
Yes	87	10.5
No	207	25.0
Not mentioned	62	7.5
Irrelevant[1]	473	
Did they describe the study location?		
Yes	659	79.5
No	51	6.1
Irrelevant[1]	119	14.4
Did they mention the incidence/prevalence of malaria?		
Yes	248	30
No	122	14.7
Irrelevant[1]	459	55.3
Did they mention the frequency of mixed infection?		
Yes	46	5.5
No	306	36.9
Irrelevant[1]	477	57.6
Final decision based on the contents of abstracts		
Suitable for meta-analysis	104	12.5
Do not have required data	657	72.2
The full text of paper is needed to make decision	68	8.3
Total	829	100

1: The abstracts were about some issues that were completely irrelevant to the questions, for example they may discuss about the prediction of malaria in future or the best health policy to control malaria in the word (appendix one).

5.2.5. Reviewing the full texts

In this stage the full text of 172 papers were reviewed and the data of eligible papers was abstracted. Out of 104 papers that were classified as suitable based upon their abstracts, 42 papers were eligible (40.4%). Out of 68 papers for which their appropriateness was not defined based upon the abstracts, 18 (26.4%) papers were eligible. The full text of one paper in

Chinese was excluded, but papers in other languages (French, Portuguese and Arabic) were assessed.

Four papers reported the frequency of infections in more than one population. These papers recruited distinct populations in different geographical locations. The data for these populations were treated independently.

The abstracted data from the full texts of eligible papers included the duration of the studies, their geographical spans and their locations, the age group of subjects, subjects' disease status (normal, febrile), the number of examined blood slides and the number of positive slides for only *P. falciparum*, *P. vivax* and for both species (Appendix 1).

Table 5-2 shows the geographical and temporal span of these studies. The minimum and maximum numbers of examined slides were 95 and 986,127 respectively (mean=23,058). Eight studies did not find any mixed infections, while one study found 782 (mean=33). Out of 62 studies, 26 (42.9%) studies sampled febrile subjects and 5 (7.9%) studies sampled children exclusively.

Table 5-2: Descriptions of eligible studies for meta-analysis

	Frequency	*Percentage*
Continent		
Asia	52	83.9
Africa	4	6.4
America	6	9.7
Spatial span		
Villages	36	58.1
District	16	25.8
Province or larger	10	16.1
Temporal span		
Month	26	41.9
Season	12	19.3
Year	5	8.1
Greater than one year	19	30.7
Age group		
Children	5	8.1
All age groups or adults	57	91.9
Samples		
Febrile	26	41.9
Normal	36	58.1
Total	62	100

5.3. Statistical methods

5.3.1. Overview of the methods

The data of every study were summarised in two by two tables and the odds ratios (ORs) between *P. falciparum* and *P. vivax* were computed.

The analysis was started by simple statistical tests ignoring any heterogeneity between studies and assessed the significance between observed and expected number of mixed infections. Then, the differences in ORs of subgroups were assessed.

In the next step, random effects meta-analysis was done to summarise the overall OR and compute its confidence interval.

In the last step, meta-regression used to assess the sources of heterogeneity in the study ORs and estimate the adjusted overall OR.

5.3.2. Simple methods

In the first step, simple statistical tests were used to check the significance between observed and expected numbers of dual infections. In this step any heterogeneity between studies was ignored. For this purpose, Pearson chi-square was used to test the significance of each study. Then, Mantel-Haenszel chi-square was used to check the discrepancies between observed and expected numbers of dual infections adjusted for the ID number of each study.

In the next step, the OR between *P. vivax* and *P. falciparum* was estimated in each study. Using ecological analysis (analysing data at the group level), the relationships between the ORs and the prevalence of malaria, the number of examined slides, the temporal and spatial span of studies, the geographical location of the studies (continent) and the age group of subjects were calculated.

5.3.3. Meta-analysis methods

Finally, the degree of heterogeneity in ORs explained by other variables was assessed. Using the random effect meta-analysis method, a summary of the OR was estimated. Then it was determined how much of the observed heterogeneity could be explained by age group, continent, endemicity, geographical size of study areas and duration of studies.

Ignoring the possibility of residual heterogeneity (i.e., between study variance) would underestimate the standard errors of the regression coefficients and thus overstate the importance of covariates.

To overcome this problem there are two main approaches *[143-145]*. The first method is an extension of weighted regression and can be applied to all types of outcomes while the second method is an extension of logistic regression and can be applied just for dichotomous outcome. Both methods estimates another extra term (τ^2) as a measure of between studies heterogeneity (Equation 5-3).

Weighted regression assumes that the log OR between *P. falciparum* and *P. vivax* (y_i) of different studies (i=1 to k) follows a Gaussian distribution. This regression needs to be weighted to take into account the precision of the estimated log ORs. The simplest model, without allowance for residual heterogeneity, is:

$$y_i \sim N(\alpha + \beta.\boldsymbol{X_i}, v^2) \qquad \text{Equation 5-1}$$

where $\beta.\boldsymbol{X_i}$ is the scalar product of β. and $\boldsymbol{X_i}$; $\boldsymbol{X_i}$ is a vector of explanatory variables in study i, β represents a vector of coefficients, α represents the log OR in the baseline of $\boldsymbol{X_i}$, and v^2 is the variance of the log OR between studies. Maximum likelihood estimates of α and β can be obtained by least squares regression of y_i on $\boldsymbol{X_i}$ with weights $w_i=1/v^2$

To incorporate residual heterogeneity (variation of true effect between studies) multiplicative factor greater than 1 could be applied to each variance, as in the following model:

$y_i \sim N(\alpha + \beta.\boldsymbol{X_i}, \varphi v^2)$ Equation 5-2

where φ is an over dispersion parameter. This multiplicative method might introduce some bias due to the dominance of large studies over small ones.

The error can be overcome by including an additive term. This involves a new term (τ^2) which represents the residual heterogeneity as follows:

$y_i \sim N(\alpha + \beta.\boldsymbol{X_i}, v^2 + \tau^2)$ Equation 5-3

Maximum likelihood method estimates α and β by regression of y_i on $\boldsymbol{X_i}$ with weights $w_i^*=1/(v^2+\tau^2)$. Thompson (1999) explained four methods to estimate τ^2: maximum likelihood, restricted maximum likelihood, moment and empirical Bayesian methods *[144]*.

Logistic regression uses binomial structure of the data. The conventional logistic regression model (without allowance for residual heterogeneity) is:

$logit(\pi_{ij}) = \gamma_i + \alpha z_{ij} + \beta x_{ij}$ Equation 5-4

where π_{ij} is the true risk of *P. falciparum* infection in group j (j=0 is for the absence of *P. vivax* and j=1 is for presence of *P. vivax* in study i which is shown by z_j as an indicator variable for the group). The other parameters are the same as previous equations. Inclusion of γ_i provides for a stratified analysis.

To correct the over-dispersion of SEs (standard errors), we could use multiplicative or additive methods. In the multiplicative method the SEs are adjusted by the following equation:

$var(y_{ij}) = \varphi n_{ij}(1-\pi_{ij})$ Equation 5-5

In this equation φ is an over-dispersion parameter, which can be estimated by dividing Pearson χ^2 by the residual degrees of freedom of the model. The variable n_{ij} is the number of subjects and y_{ij} is the numbers of events in the j group of study i. This method has the same problem as the multiplicative model discussed in the previous Section.

In contrast, the additive model allows the variability between studies by multi-level method:

logit $(\pi_{ij}) = \gamma_i + \alpha z_{ij} + \beta x_{ij} + \alpha_i z_j$ Equation 5-6

where α_i is a random effect , drawn from a Gaussian distribution with mean 0 and variance τ^2, and expresses the way in which the log-odds ratio of study i deviates from the value expected from the other explanatory variables. Using restricted iterative generalised least squares,τ^2 can be estimated *[144]*.

The first method, weighted regression, has been used more commonly in meta-analysis papers. Also, it is more easily applied in Stata. Therefore, in this analysis, the OR was modelled by the weighted regression method. Also the additive term was utilised to estimate the effect of explanatory variables and the residual heterogeneity.

Using the "*metareg*" command of Stata, the probability of dual infections was predicted by entering in the probabilities of single infections. In order to compare the residual heterogeneity (τ^2) in random effect models with and without explanatory variables, the moment method was used. One reason

for this choice is that maximum likelihood and even restricted maximum likelihood methods introduced a downward bias in estimating of τ^2 *[144]* which is also observed in this study (the estimated τ^2 values were less than 0.0005).

5.4. Results

The data of 62 studies were included in this meta-analysis. They were published between 1975 and 2001. The smallest and largest studies examined 95 and 206,997 blood slides respectively.

Eight studies reported zero mixed slides. To utilise their data in logarithm scale, 0.5 was added to all cells.

Table 5-3: The risk ratio of *P. falciparum* in *P. vivax* positive versus *P. vivax* negative group; studies are sorted by their publication year

First author; year	Odds Ratio (95%CI)	First author; year	Odds Ratio (95%CI)
J. H. Cross;1975*[146]*	1.53(0.37-6.41)	M. Giboda;1992*[147]*	0.16(0.08-0.33)
M. Maffi;1975*[148]*	4.43(1.31-15.04)	R. L. Anthony;1992*[149]*	0.03(0-0.44)
R. Rajagopal;1976*[150]*	1.94(0.12-32.75)	Syafruddin;1992*[151]*	0.28(0.16-0.5)
J. H. Cross;1976*[152]*	0.43(0.16-1.17)	T. Adak;1994*[153]*	1.2(0.26-5.6)
L. L. Smrkovski;1982*[154]*	10.92(3.39-35.2)	P. Dutta;1994*[155]*	0.5(0.12-2.02)
J. Cattani;1983*[156]*	0.44(0.05-3.57)	S. Rafi;1994*[157]*	0.86(0.24-3.1)
J. Hii;1985*[158]*	1.14(0.67-1.94)	S. Das;1994*[159]*	0.71(0.1-5.09)
J. Hii;1985*[158]*	5.16(1-26.72)	Y. Mizushima;1994*[160]*	0.1(0.03-0.34)
M. De Arruda;1986*[161]*	0.22(0.03-1.61)	P. Gautret;1995*[162]*	0.84(0.56-1.28)
J. A. Cattani;1986*[163]*	2.91(0.16-51.4)	S. K. Ghosh;1995*[164]*	1.75(0.1-31.35)
B. L. Verma;1986*[165]*	0.46(0.4-0.53)	R. Dietze;1995*[166]*	0.39(0.22-0.72)
A. E. Beljaev;1987*[167]*	0.15(0.07-0.35)	P. Dutta;1995*[168]*	0.54(0.12-2.33)
G. T. Strickland;1987*[169]*	0.3(0.17-0.51)	J. Y. Uchida;1995*[170]*	0.16(0.01-2.75)
G. T. Strickland;1987*[169]*	1.17(0.74-1.85)	J. B. Sherchand;1995*[171]*	0.3(0.13-0.69)
G. T. Strickland;1987*[169]*	0.6(0.31-1.14)	N. G. Das;1997*[172]*	0.19(0.17-0.21)
G. T. Strickland;1988*[173]*	1.73(1.04-2.86)	F. W. Hombhanje;1997*[174]*	0.92(0.85-0.99)
N. Singh;1989*[175]*	0.79(0.57-1.08)	V. Y. Belizario;1997*[176]*	0.07(0.01-0.56)
N. Singh;1989*[175]*	0.94(0.75-1.17)	V. Y. Belizario;1997*[176]*	9.56(4.5-20.32)
P. Dutta;1989*[177]*	0.13(0.03-0.58)	V. Y. Belizario;1997*[176]*	2.58(0.35-18.94)
H. Itokawa;1989*[178]*	0.97(0.69-1.37)	T. Seboxa;1997*[179]*	1.68(0.71-3.98)
S. K. Ghosh;1989*[180]*	0.05(0.03-0.07)	W. P. Carney;1977*[181]*	0.95(0.13-7.04)
P. M. Graves;1989*[182]*	0.1(0.04-0.3)	B. Mandal;1998*[183]*	0.02(0.01-0.04)
E. Fox;1989*[141]*	0.16(0.01-2.73)	H. Joshi;1998*[184]*	0.14(0.03-0.59)
L. K. Das;1989*[185]*	0.83(0.33-2.07)	L. M. A. Camarg;1999*[186]*	0.34(0.08-1.55)
P. Dutta;1990*[187]*	0.24(0.01-4.4)	Singh Neeru;2000*[188]*	2(0.93-4.31)
P. K. Rajagopal;1990*[189]*	0.32(0.04-2.59)	S. Hozhabri;2000*[190]*	3.49(1.1-11.09)
F. Nosten;1991*[191]*	2.75(1.62-4.64)	R. K. Mehlotra;2000*[192]*	0.17(0.11-0.27)
S. Subramanian;1991*[193]*	3.11(1.11-8.72)	H. C. Srivastav;2000*[194]*	1.36(0.98-1.88)
M. da L. R. Moitinho;1991*[195]*	0.34(0.19-0.61)	J. Pinto;2000*[196]*	1.24(0.07-22.18)
D. M. Gordon;1991*[197]*	0.31(0.15-0.62)	M. H. Roper;2000*[198]*	0.65(0.24-1.76)
P. Dutta;1991*[199]*	3.09(1.42-6.73)	N. Singh;2001*[200]*	0.51(0.18-1.39)

The overall OR was 0.6 (Mantel-Haenszel chi^2= 1205.32, df=1, p<0.0001, ignoring heterogeneity: 95% CI: 0.46-0.79); the minimum and maximum ORs were 0.02 and 10.9 respectively (Table 5-3). Out of 62 studies, 41 (66.1%) showed OR less than one; 20 of them (48.8%) were statistically significant. Among those studies with OR greater than one, 8 (38.1%) had significant p-value.

There were significant associations between the ORs and frequencies of *Plasmodium* spp. Pearson correlation coefficients between the ORs and the frequency of *P. falciparum, P. vivax* and all species were -0.33, -0.34 and -0.44 respectively (p<0.0001). These results show that the OR had negative associations with the frequencies of species i.e., ORs were greater in low endemic areas. However, the Pearson correlation coefficient between the number of examined slides and the OR was -0.05 (p=0.76). In other words, the OR was not influenced by the size of studies.

Potential explanatory variables in these analyses were age group (children or mixed), continent, study group (febrile or normal), temporal and spatial spans of studies and frequencies of *P. falciparum* and *P. vivax* infections among examined blood slides.

There were highly significant heterogeneities among the ORs even in subsets of studies (Table 5-4). Moment method (Section 5.3) was used to quantify the heterogeneity between ORs (τ^2). The overall τ^2 was 0.91. It was even greater than this value in most subsets of studies classified by explanatory variables. However, it was considerably lower in studies with low *P. vivax* frequency, high *P. falciparum* frequency, in studies with short

temporal or wide spatial spans and among normal subjects and in studies in South America.

The summary OR in South American studies was significantly lower than those in other continents. It was 0.21, 0.62 and 1.76 in South America, Asia and Africa respectively. Although the difference between the ORs in Africa and Asia was considerable, it was not statistically significant.

The summary OR in studies that recruited only children was greater than that in studies which recruited all age groups (1.28 versus 0.56). The confidence interval around OR in children was very wide (0.31-6.08) which was due to a few available studies (five) and considerable residual heterogeneity (τ^2=2.14). Nonetheless, it is compatible with the results of many studies which reported higher risk of mixed infections among children.

The summary OR of studies in normal subjects was around twice of that in febrile subjects (0.9 versus 0.35; p=0.04). In other words, the risk of mixed infection in febrile subjects was less than that normal subject.

There were not any obvious trends in the summary ORs classified by temporal and spatial size of study. The maximum ORs were observed in middle size studies; ORs were lower in long studies (one year or longer: 0.39) or studies in wide areas (larger than a district: 0.49).

Lower summary ORs were observed in studies with higher frequencies of infections. The OR in studies with frequency of infection 30% or more was 0.32 (95% CI: 0.22-0.47); while the corresponding OR in studies with frequency less than 15% was 2.51 (95% CI: 1.66-3.8). Similar descending trends in the ORs were observed classified by the frequencies of *P. falciparum* and *P. vivax* infections.

Lower summary OR in South American studies may be explained by the higher *P. vivax* frequencies. The average of *P. vivax* frequency in South American studies was 20.2% (SD=10.6); the corresponding values in African and Asian studies were 6.7 (SD=7.3) and 6.5 (SD=4.7) respectively. Therefore, lower ORs in South America may be explained by negative association between the frequency of *P. vivax* infection and the OR.

The difference between summary ORs in febrile and normal subjects can also be explained by differences in the frequencies of *Plasmodium* spp. The average of frequencies of infections in febrile and normal subjects were 30.2% (SD=19.7) and 21.0% (SD=13.5) respectively. Most of this difference was due to the difference in the frequencies of *P. falciparum* infection (22.7% versus 15.1%).

Furthermore, the frequencies of infections in studies among children was lower than those in mixed age group studies (7.9%, SD=5.7 versus 19.2%, SD=15.6). Therefore, the greater summary OR in the former group can be due to the lower frequencies of infections.

Based on the above findings, it seems that the frequencies of infections are the only main explanatory variables in describing the patterns of the ORs in subsets of studies.

Table 5-4: The odds ratio of *P. vivax* as risk factor of *P. falciparum* classified by continent, age group, study subjects, temporal and spatial span and the frequencies of species among examined slides

Subgroup (number of studies)	*Odds ratio (95%CI)*	*P-value of heterogeneity*	*Tau square*
Continent			
Asia (52)	0.62(0.46-0.83)	<0.0001	0.8
South America (6)	0.21(0.16-0.26)	0.0013	0.59
Africa(4)	1.76(0.47-6.6)	<0.0001	1.68
Age group			
Children(5)	1.38(0.31-6.08)	<0.0001	2.14
Mixed(57)	0.56(0.43-0.75)	<0.0001	0.82
Subjects			
Normal(36)	0.9(0.65-1.24)	<0.0001	0.68
Febrile(26)	0.35(0.21-0.58)	<0.0001	1.33
Spatial span			
A few villages(36)	0.5(0.33-0.75)	<0.0001	1.14
District(16)	0.99(0.591-1.63)	<0.0001	0.73
Larger than a district(10)	0.49(0.3-0.82)	<0.0001	0.43
Temporal Span			
Month(26)	0.81(0.56-1.17)	<0.0001	0.55
Season(12)	0.97(0.52-1.79)	<0.0001	0.66
Year or longer(24)	0.39(0.26-0.6)	<0.0001	0.88
P. falciparum risk (%)			
<10(23)	1.06(0.54-2.1)	<0.0001	5.15
10-14.99(10)	0.75(0.42-1.35)	<0.0001	0.6
=15(29)	0.4(0.28-0.57)	<0.0001	0.77
P. vivax risk (%)			
<5(27)	1.43(0.98-2.1)	<0.0001	0.5
5-9.99(18)	0.49(0.32-0.75)	<0.0001	0.6
=10(17)	0.25(0.13-0.5)	<0.0001	1.72
Both species risk (%)			
<15(18)	2.51(1.66-3.8)	0.0021	0.36
15-29.99(22)	0.5(0.36-0.7)	<0.0001	0.37
=30(22)	0.32(0.22-0.47)	<0.0001	0.62
All studies (62)	0.6(0.46-0.8)	<0.0001	0.91

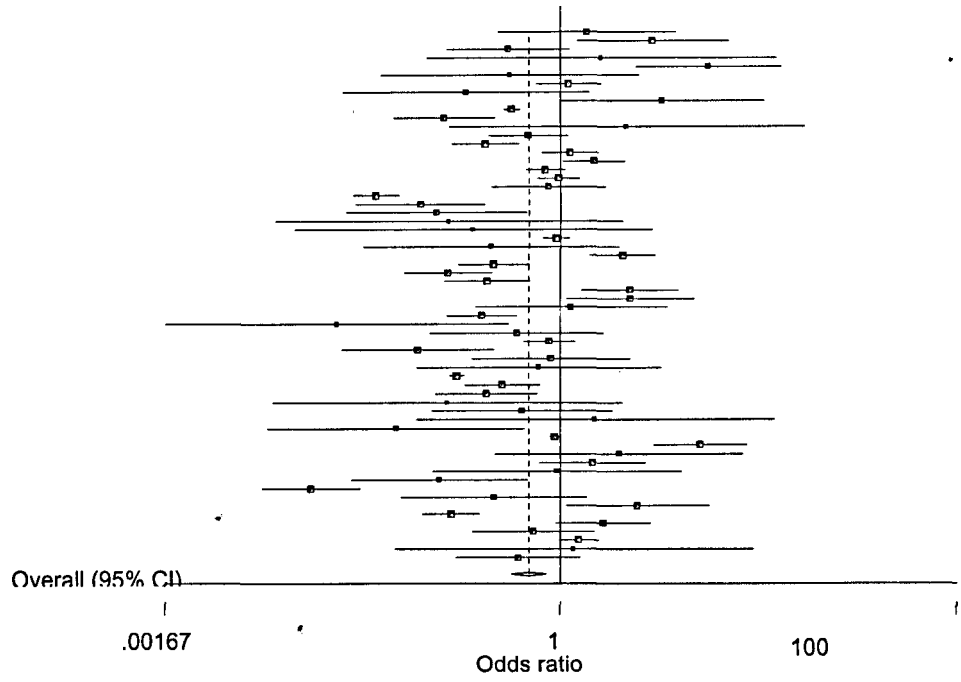

Figure 5-1: Forest plot of the OR between *P. vivax* and *P. falciparum* using random effect model to estimate the confidence interval around the summary OR (diamond figure)

The results of regression models show that only *P. falciparum* and *P. vivax* frequencies explained some part of the residual heterogeneity (Table 5-5). Although in these models temporal span of studies showed negative association with the OR ($p<0.03$), it did not change the τ^2 considerably (model with the frequency of all species: $\tau^2= 0.73$, model with both the frequency of all species and temporal span of studies: $\tau^2=0.72$). None of the other potential explanatory variables were significant and they mostly inflated τ^2 (0.91 versus 1.18). However, both *P. falciparum* and *P. vivax* frequencies dropped τ^2; accounting for their joint effects, τ^2 was 0.73. Therefore, it seems that the frequencies of infections were the only

important predictor on the risk of mixed infection which explained a considerable part of residual variations (more explanation in Section 5.5.6).

Compatible with the pattern of ORs in subsets (Table 5-4), the modelling results show that the OR was greater in studies with low frequency of infection. Also, there were a negative association between the temporal span of studies and the OR between *P. falciparum* and *P. vivax*.

Table 5-5: Impacts of potential explanatory variables on the residual heterogeneities between the ORs

Subgroup	*Main finding*	*Tau square*
Model 1: no explanatory variable OR	0.6(0.46-0.8)	0.91
Model2: explanatory variables were age group, subjects (febrile or normal), spatial and temporal span of studies and continent	only fever and age group were significant	1.18
Model3: explanatory variables were the variables in model 2 and the frequencies of *P. falciparum* and *P. vivax*	Only the frequencies of *P. falciparum* and *P. vivax* and temporal span of studies had significant p-value; all of them had negative associations	0.96
Model4: explanatory variables were only the frequencies *P. falciparum* and *P. vivax* and temporal span of studies	Both *P. falciparum* and *P. vivax* had significant p-value	0.77
Model5: the only explanatory variable was the frequencies of all species (all *Plasmodium* species considered together) and temporal span of studies	Both variables had significant negative associations	0.72

5.5. Discussion

This section reviews the main findings and then explains the differences between summary ORs. It also compares the meta-analysis results with others' findings.

This systematic literature review only included published papers. Estimation of the mixed infections frequency was not the main objective of almost all of the eligible papers. Therefore, there is not a convincing reason to believe that the chance of their publications had any correlation with mixed infections frequencies. Based on this reason, it can be expected for that publication bias not to be an important issue in this meta-analysis.

5.5.1. Description of main findings

The systematic review of the literature found 62 eligible studies between 1975 and 2001. Surveillance data were not included in this meta-analysis to minimise the possible error of misclassification of mixed infections, because research studies usually examine blood slides more accurately.

The OR of *P. vivax* and *P. falciparum* infection was computed based on the data from these studies. The OR varied over a wide range from 0.02 to 10.9. The summary OR was 0.6 (random effect 95% CI: 0.46-0.79). It means that in overall, one species infection decreases the risk of the other species.

The ORs in studies that recruited children or subjects without fever were greater. Also, the summary OR of studies from South America was significantly lower than those in studies from Asia or Africa. However, these

differences could be explained based on the differences in the frequencies of infections.

The above explanation was also supported with the meta-regression results (Table 5-5). Both age group of subjects and presence of fever were significant (model 2) but their significance disappeared by adding the frequencies of *P. falciparum* and *P. vivax* (model 3).

In addition, the modelling shows that a considerable part of the heterogeneity between ORs can be explained by the frequencies of *Plasmodium* spp among the examined blood slides. The residual heterogeneity between ORs of all studies was 0.91; having taking into account of the frequency of all infections, it dropped to 0.72. Also, the temporal span of studies had significant negative association with the OR, but it did not change Tau-square considerably. None of the other potential explanatory variables were significant and they mostly inflated Tau-square.

5.5.2. Interaction between *Plasmodium* spp

Sections 4.3 and 4.4 explain the possible explanations for the positive and negative interactions between *Plasmodium* spp. The following sections relate these explanations with the main results of the meta-analysis.

5.5.3. Temporal and spatial span of studies

High risk group for one species may not have high infection risk for the other species simultaneously. Therefore, discrepancies between temporal

and spatial variations of *Plasmodium* spp risks are one of the possible explanations for negative interaction between species.

It might be expected that the discrepancies to be greater in larger studies which observed people in longer period or wider area.

This meta-analysis checked the effect of temporal and spatial spans of studies. No trend in ORs was observed based on the spatial span of studies. However, temporal span had a negative association. It means that generally, the OR is longer studies were smaller.

Therefore, it may imply that the spatial span of studies is not accurate measures of the discrepancies in infection risks. However, the temporal span of studies play a role on the observe interaction between species.

5.5.4. Geographical location

Most of the included studies were from Asian countries. This is because *P. vivax* is generally common and many studies have been carried out to assess the epidemiology of *P. vivax* in this continent.

Howard et al. (2001) assessed the risk of infections with multiple parasite species and a log-linear method was presented for analysing data from multiple communities and testing whether the associations in different communities were equal *[125]*. He did not systematically review of the literature and assessed the associations between many parasitic infections. He reported largely negative associations between *Plasmodium* spp in Asian countries, however, positive associations were found in Tanzania, Papua New Guinea and the USA *[125]*.

The geographical pattern of association in this meta-analysis was not compatible with Howard's findings. It seems that geographical location is not an independent factor; the results of this meta-analysis show that lower OR in South American studies may be due to higher frequency of infections.

Therefore, the differences between Howard's findings and the results of this meta-analysis could be only due to the differences in the frequencies of infections in the recruited studies.

5.5.5. Effect of fever and age

The crude ORs in febrile subjects were less than those in normal subjects. Acquired immunity against malaria has negative correlation with fever *[80-83]*. Therefore, it may be implied that mixed infection was less common in those who had not acquired immunity.

The above conclusion is not compatible with the differences between crude ORs in studies that recruited children and mixed age groups. Children usually have less immunity while their summary OR was higher and this showed that the risk of mixed infection in the young age group was higher. This finding is also supported with the results of many other studies.

This contradiction may be explained by the fact that this is a meta-analysis and it assessed the risk of mixed infections in studies not individuals. Therefore, it should be paid enough attention to some confounding factors such as the frequency of infections.

In the regression models, the frequencies of infections were the only significant variables and explained a considerable part of the residual

heterogeneity. Having taken into account of the frequencies, the effects of both variables were not significant (Table 5-5).

There were also clear differences between the frequencies of infections in subsets of studies which recruited in febrile and normal subjects, and studies which surveyed children and all age groups. The OR was lower in studies with high infection frequencies and the effects of age and fever could be explained based on the differences in frequencies of infections.

5.5.6. Frequencies of infections

Higher frequency of infection in blood slides particularly in normal subjects was correlated with the prevalence (or incidence) of infections. Therefore, it can be implied that the interaction is negatively correlated with the risk of infections in the population.

McKenzie and Bossert (1997) showed that high overall prevalence of infection is associated with significant deficits of dual infections based on the product of individual species prevalences *[126]*. He explained this finding based on biological interactions between species but He did not discuss about any specific pathways.

Some of findings in McKenzie's study was similar to the findings in this meta-analysis although this analysis was based on OR not the product of prevalences, i.e., risk ratio. In addition, this analysis assesses the effects of temporal and spatial span of studies and other possible explanatory variables.

Regarding to the possible explanation between the interaction and the frequencies of species, the differences between group and individual data should be noted. Both of these two studies checked the associations in group data. Although, it is not possible to rule out the possible effects of biological interactions, it should be taking into account the other possible confounder factors which may distort relationships in group data (cross level confounding effect).

One of these possible variables might be the heterogeneity in infection risks in a population. Chapter 6 illustrated the positive impact of heterogeneity on the interaction between species. However, it is very difficult to measure the heterogeneities in infection risks within populations even in epidemiological studies. Therefore, there are not any explicit data to quantify its association with the infection prevalences.

The acquired immunity is an alternative rational explanation. In highly endemic area people usually have stronger immunity which may protect them against mixed infections.

5.5.7. Final conclusion

The overall OR between *P. vivax* as risk factor of *P. falciparum* was less than one and shows negative interaction between these two species.

It seems that the prevalence of infections in the populations has negative association with the OR and decreases the residual heterogeneity between studies ORs. This negative association may be due to either biological interaction between species *[126]* or the effect of heterogeneity in exposure risks among the population which may confound the analysis of group data.

The crude effects of other factors such as age group, continent and presence of fever can be explained by the frequency of infections.

It was also shown that the temporal span of studies had negative association with the interaction between species. Therefore, some part of negative associations in long studies may be explained based on the discrepancies between the temporal variations of species.

CHAPTER 6

6. Modelling the heterogeneity effect

In this chapter, the effect of heterogeneity in infection risks on the overall interaction between species is assessed. The possible explanations of the positive interactions between species are discussed. A number of models which assess the impact of the heterogeneity in infection risks on the interaction between species are then presented.

6.1. Definition of the infection risk heterogeneity

Positive associations between species mean that a subgroup of people, in terms of time or space, has higher infection risks for all species, i.e., heterogeneity in infection risks within the population. For example, suppose that children get infections more than adults. Then children could be a subgroup which 'has higher infection risks for all species' because they have higher risk than adults. However, within children, there may be no species interaction (i.e., there may be sampling independence). So at the population level there might appear to be an interaction, which would disappear on stratifying by age. In this thesis the above explanation is called "heterogeneity hypothesis".

Section 4.3 presents two main explanations of the positive associations between species, similarity in transmission routes and higher susceptibility of a subgroup of people, which are not mutually exclusive.

These two explanations can be easily linked to the heterogeneity hypothesis. The first explanation, similarity in transmission route, describes the heterogeneity in exposure risks; i.e., a subgroup of people has higher exposure risks to all species, and the second explanation describes the heterogeneity in susceptibility, i.e., a subgroup of people is more susceptible to infections by all species. Both of the heterogeneities in exposure risks and susceptibilities reflect different aspects of the heterogeneity in infection risks.

6.2. Confounding effect of the heterogeneity

A confounder is a variable which distorts the observed association between two variables. In other words, a confounder decreases or increases the magnitude of the observed association between explanatory and outcome variables.

A confounder must have independent associations with both explanatory and outcome variables and not be on the causal pathway.

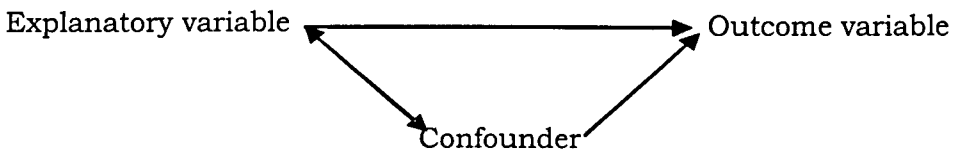

Figure 6-1: Schematic relationship between the confounder, explanatory and outcome variables

The individual infection risks can be considered as a confounder on the interaction between species if it varies within population. Considering one species, e.g. *P. falciparum,* as a risk factor for other species, the infection risk has positive associations with both *P. falciparum* (explanatory variable) and the other species (outcome variables). Therefore, a crude analysis in the whole population might show positive associations artificially, i.e., OR greater than one, as the result of the confounding effect of the heterogeneity in infection risks.

Figure 6-2 is a schematic diagram which illustrates the mechanism of the heterogeneity effect on the crude OR. It divides population into high and low risk groups. The ORs between species in both subgroups are one, meaning no association between species in the subgroups. However, the overall OR is more than one, due to the confounding effect of the heterogeneity in infection risks.

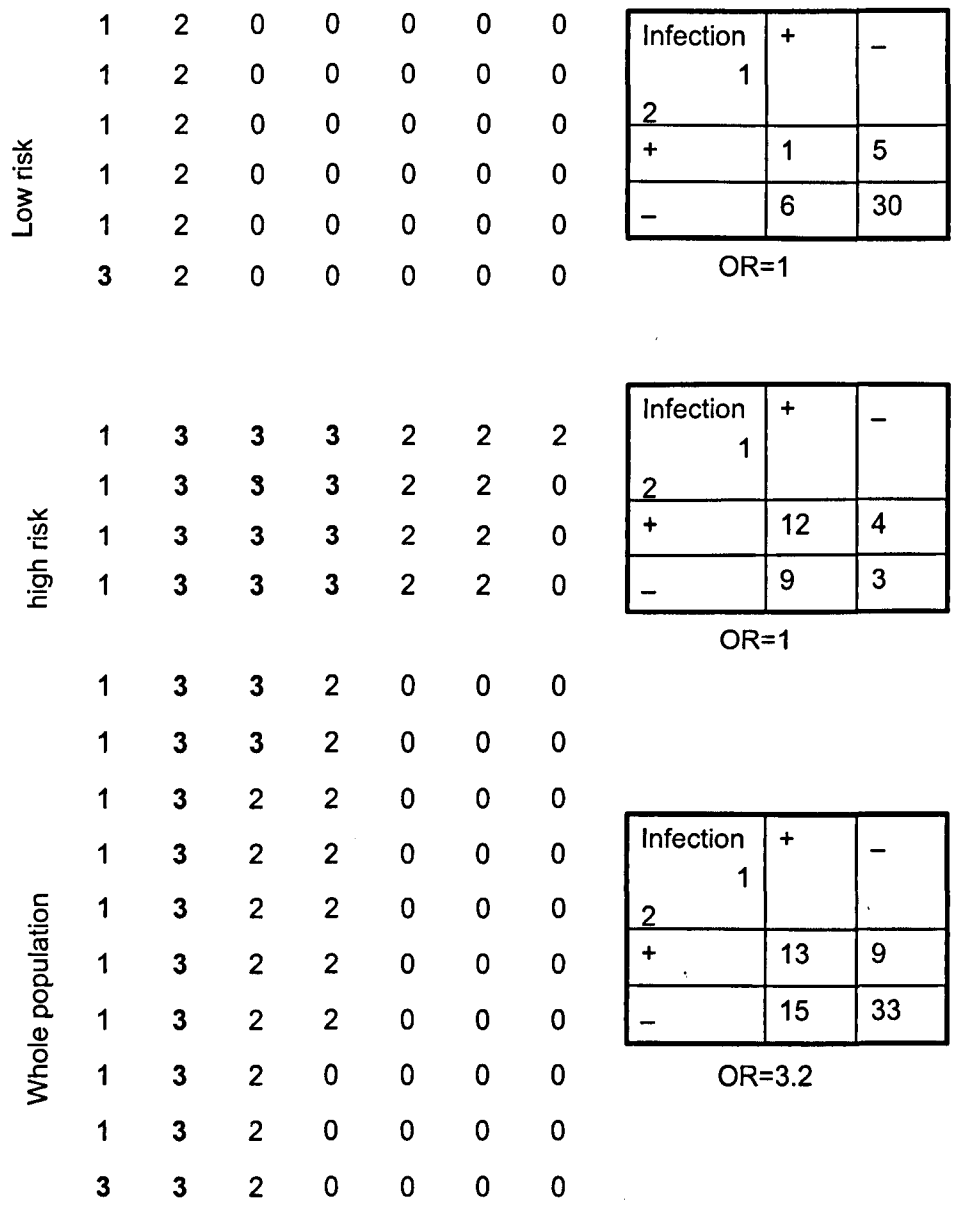

Figure 6-2: Schematic diagram illustrating the confounding effect of the heterogeneity in infection risks on the crude association between species
(0, 1, 2, and **3** represent no infection; infection with species number one; species number 2; and mixed infection respectively)

Nevertheless, based on the above suggestion, a positive association should exist between the risks of all infections; i.e., the high-risk people should be prone to all types of *Plasmodium* spp infections.

Strong positive associations between species were reported as high as OR around ten (Table 5-3). The following section assesses if heterogeneity in infection risks can, by itself, explain these very strong positive associations.

6.3. Heterogeneity models

6.3.1. Description

This section models the impact of heterogeneity in infection risks among people on the overall OR, and assesses whether the heterogeneity on its own can explain ORs as high as those observed in some of the studies in the meta-analysis chapter (Table 5-3).

Simple models were created using Microsoft Excel 2000. The models assumed that the population consisted of two layers, low and high-risk groups, with no real associations between infections in each layer (stratum specific OR=1).

In the main models, the infection prevalences in the low-risk group were set to 0.17 and 0.06 (the minimum frequency of *P. falciparum* and *P. malariae* infections among villages and age groups in the Garki data, see next chapter). Nonetheless, the model sensitivity to these frequencies was assessed.

The infection frequencies in high-risk group were computed as the product of "k_i" (i=1, 2 for the first and second species respectively) and infection frequencies in low-risk group. In other words, k_i was the risk ratio of infection i in high-risk versus low-risk groups, and varied between one up to its maximum valid values of 5.9 and 16.6 for the first and second infections respectively (the maximum risk of infections in the high-risk group could not be more than one; therefore, k_1=1/0.17=5.88 and k_2=1/0.06=16.6). The model simulated the effects of k_i, and the ratio between the numbers of people in the high and low risk groups, ($m=N1/N0$) on the overall OR.

6.3.2. Model structures

k_i indicates the ratio of infection risk in high versus low risk group ($i=1$ and 2 stands for species one and two). Also, subscript j stands for risk group, $j=0$ and 1 show low and high risk groups.

The following table shows the numbers of positives and negatives in group j for species one and two.

Table 6-1: Cross tabulation between species one and two

species two / species one	+	-	Total
+	$A_{.j}$	$B_{.j}$	N_{1j}
-	$C_{.j}$	$D_{.j}$	$N_{.j}$-N_{1j}
Total	N_{2j}	$N_{.j}$-N_{2j}	$N_{.j}$

Assuming $N_{.1}=N_{.2}$, i.e., equal size of low and high risk groups, it could write

$N_{i2}=N_{i1} \cdot k_i$

In other words, the number of infected people by species *i* in high risk group is the product of the corresponding infected people in low risk group and its risk ratio.

$OR_{.0}$ and $OR_{.1}$ are the odds ratios between two species in the low and high risk groups, which are one based on the independence assumption.

$OR_{.t}$ is the odds ratio between two species in the whole population, and it would be estimated by $OR_{.0}$, $OR_{.1}$ and k_i

In the low risk group ($j=0$):

$A_{.0}=N_{10}-B_{.0}$

$D_{.0}=N_{.0}-N_{20}-B_{.j}$

$C_{.0}=N_{20}-N_{10}+B_{.j}$

and

$$OR_{.0}=1$$

$$[(N_{10}-B_{.0}) \cdot (N_{.0}-N_{20}-B_{.0})]/ [B_{.0} \cdot (N_{20}-N_{10}+B_{.0})] = 1$$

$$(N_{10}-B_{.0}) \cdot (N_{.0}-N_{20}-B_{.0}) = B_{.0} \cdot (N_{20}-N_{10}+B_{.0})$$

Hence:

$B_{.0} = (N_{10}N_{.0}-N_{10}N_{20}) / N_{.0}$

$A_{.0} = (N_{10}N_{20}) / N_{.0}$

$C_{.0} = (N_{.0}P_{20}-P_{10}P_{20}) / N_{.0}$

$D_{.0} = [(N_{.0}-P_{10}) \cdot (N_{.0}-P_{20})]/ N_{.0}$

Having assumed $k_1=k_2=k$ and $N_{.0}=N_{.1}$

$N_{11}=k \cdot N_{10}$

$N_{21}=k \cdot N_{20}$

Hence:

$B_{.1} = (N_{10}N_{.1}k - N_{10}N_{20}k) / N_{.1}$

$A_{.1} = (N_{10}N_{20}k) / N_{.1}$

$C_{.1} = (N_{.1}N_{20}k - N_{10}N_{20}k) / N_{.1}$

$D_{.1} = [(N_{.1}-N_{10}k)*(N_{.1}-N_{20}k)] / N_{.1}$

In the whole population

$A_t = A_{.0}+A_{.1} = [(N_{10}N_{20}) / N_{.t}] + [(N_{10}N_{20}RR_2) / N_{.t}]$

$B_t = B_{.0}+B_{.1} = [(N_{10}N_{.t}-N_{10}N_{20}) / N_{.t}] + [(N_{10}N_{.t}k - N_{10}N_{20}k) / N_{.t}]$

$C_t = C_{.0}+C_{.1} = [(N_{.t}N_{20}k - N_{10}N_{20}k) / N_{.t}] + [(N_{.t}N_{20} - N_{10}N_{20}) / N_{.t}]$

$D_t = D_{.0}+D_{.1} = [[(N_{.t}-N_{10}) \cdot (N_{.t}-N_{20})] / N_{.t}] + [[(N_{.t}-N_{10}k) \cdot (N_{.t}-N_{20}k)] / N_{.t}]$

Hence:

$$OR_{.t} = \frac{2N_{.t}^2 + N_1N_2(1+k) - (N_{.1}+N_{.2})(1+N_{.1}.N_{.t})}{N_{.1}N_{.2}(-N_{.t}(k+1)(N_{.1}N_{.2})) + N_{.t}(1+k)^2 + N_{.1}N_{.2}(1+k)^2}$$

The above formula is too complicate to illustrate the relationships of $OR_{.t}$ with k and the ratio of $N_{.1}/N_{.2}$.Therefore, more explanation is presented based on the following graphs.

6.3.3. Results

There is an exponential relationship between the overall OR and the heterogeneity in the whole population. Figure 6-3 plots the effect of k_1 and k_2 on the overall OR in the whole population. In this graph, k_1 and k_2 are varying between one, up to their maximum possible values (5.8 and 16.6 for k_1 and k_2 respectively). The model also assumed that the low and high risk group populations are equal in size; i.e., their ratio (m) is one. The overall OR is very close to one when k_i is less than three; while, it increases steeply when k_i is greater than five. It implies that the confounding effect of the heterogeneity is considerable when the ratio of the risks of infections in low and high risk groups is substantial.

The overall OR is much greater (around 100) if the model allows that k_1 and k_2 take their maximum values. In other words, if the maximum values for k_1 and k_2 are estimated independently based on the prevalences of two species in the low risk group, the overall OR will increase by greatly.

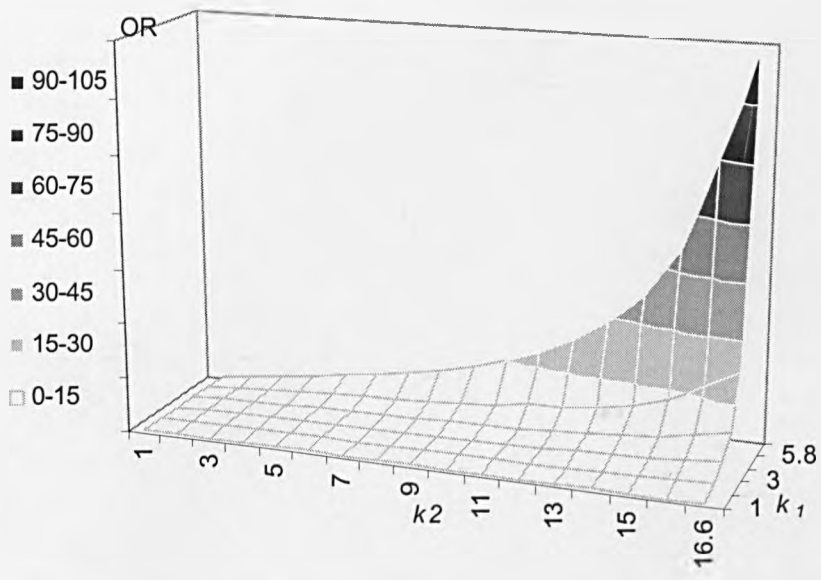

Figure 6-3: The effect of the heterogeneity in exposure risks (k_1 and k_2 in x and y axes respectively) on the overall OR (z axis)

The overall OR is not so sensitive to the ratio of populations in high versus low risk groups (m). Figure 6-4 shows the relationship between the maximum overall odds ratio and m. The graph peak is at m equal one. However the overall OR is greater than four when m is around 0.2 and greater than five when m is around five. Hence, the confounding effect of the heterogeneity is considerable and its impact is not so sensitive to the ratio of high and low risk populations in a wide range of m between 0.2 and 5.

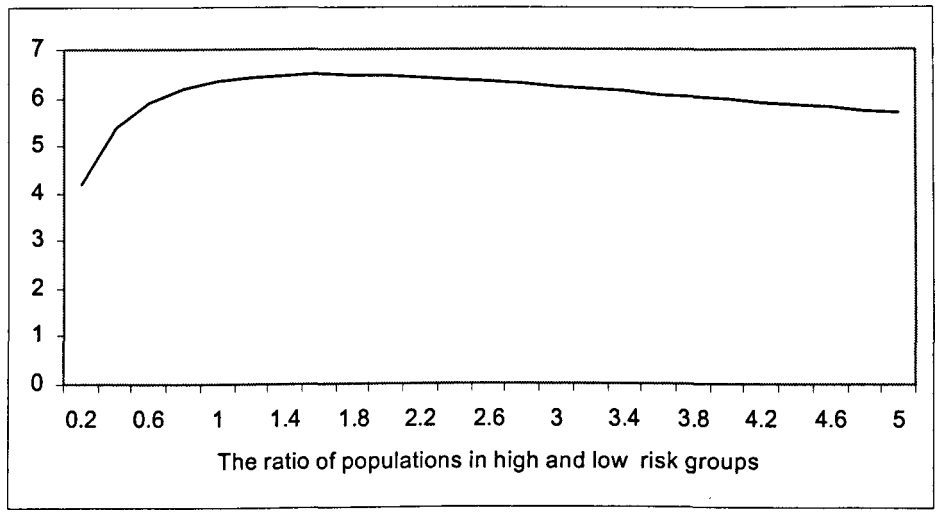

Figure 6-4: The effect of the ratio of populations in high and low risk groups (x axis) on the maximum overall OR (y axis)

In addition, the maximum possible OR is observed when the species prevalences are equal. In addition, there is a negative association between the species prevalences in the low risk group and the maximum possible OR. In other words, the possible maximum OR is greater when the prevalences are lower. That is, the lower the prevalence in the low risk group the greater the maximum possible heterogeneity in infection risks (i.e., the ratio between 100% prevalence in the high risk group, and the prevalence in the low risk group). Hence, the maximum possible heterogeneity in infection risks between low and high risk groups decreases when the prevalences in the lower risk group increases.

6.4. Discussion

Very strong positive associations between species have been reported in some epidemiological studies (Table 5-3). The models in this chapter

assessed whether the proposed reasons in Section 6.2 can explain these very strong positive associations.

It was explained how heterogeneity in infection risk can confound the association between species. In addition, its confounding effect has been quantified in simple models.

The degree of heterogeneity, the prevalences of infections in the low risk group, and the size of the low and high risk groups affect the confounding effect. The models illustrated that the confounding effect was not considerable when the heterogeneity in infection risk is low; i.e., the ratio of infection risk in high versus low risk group was less than three. However, the confounding effect rose exponentially if the heterogeneity increased and theoretically it could inflate the OR up to one hundred.

Nonetheless, the confounding effect of the heterogeneity was relatively robust to the population sizes in low and high risk groups. In other words, the confounding effect was substantial in a wide range of the ratio between population sizes of the high and low risk groups.

The confounding effect was reduced when the prevalences in the low risk group increased. The negative association was solely due to the effect of prevalences in low risk group on the maximum possible heterogeneity. Since the maximum prevalence in high risk group was one, by increasing the prevalences in low risk groups, the ratio of prevalences in low and high risk groups; i.e., the heterogeneity in infection risks, decreased.

In conclusion, based upon the model results, it can be implied that merely the heterogeneity in infection risks could explain the observed strong positive relationship between species in some epidemiological studies.

Therefore, there is not necessarily any need to explain the positive associations based on more complicated biological pathways within the human body such as the role of mixed infections on the risk of symptoms, or the degree and duration of parasitaemia. It should be added that most of the findings in molecular or even epidemiological studies do not support these pathways.

6.5. Limitations

This model divides the population into two groups, which may not be an appropriate assumption, as the infection risks have a continuum from zero up to one within a population. However, for the purpose of this analysis, it seems reasonable to believe that heterogeneity in infection risks might explain observed ORs as great as ten which has been reported in some of epidemiological studies.

CHAPTER 7

7. Interaction between species in the Garki data

7.1. Introduction

This chapter explores the interactions between *Plasmodium* species in the Garki data *[129]*, and seeks to explain the differences that are commonly found between the findings of cross-sectional and longitudinal studies (Section 7.4).

Cross-sectional and longitudinal studies report disparate findings on the frequencies of mixed species infections. Most prevalence surveys find fewer mixed-species infections than would be expected based on the product of the frequencies of individual species. This suggests that one parasite may be excluding another or suppressing the secondary species' parasitaemia to undetectable levels *[6,133]*. In contrast, only a few longitudinal studies have explored this issue. Nevertheless, a longitudinal study conducted by Bruce et al. (2000) reported that the frequency of mixed infection was very close to the number that would be expected assuming no species interaction was occurring *[134]*.

Based on these seeming contradictions, the differences between cross-sectional and longitudinal analysis must be assessed more extensively to explain these findings and test hypotheses more appropriately.

The Garki data is an extremely large, well established malarial dataset from Nigeria (Section 7.4) and even today, remains one of the best sources to study many aspects of malaria. Up until now, the published results of the Garki data had not thoroughly explored the interactions between *Plasmodium* species, and that too had only approached this issue cross-sectionally using very simple methods. Exploring the Garki data both cross-sectionally and longitudinally and assessing the effects of seasonality, age and other important variables may clarify if study design provides differing results about species interaction, and provide more evidence about any possible patterns of these interactions.

This chapter explores different aspects of the interaction between *Plasmodium* species using an extensive, well-established repository of malaria data, the Garki project. It explains the findings based on short and long term acquired anti-parasite immunity, cross-immunity and the suppression effect of one *Plasmodium* species on the blood density of another species. In this study, short-term effects are any effects which protect subjects for months to at maximum a few years; while long-term effects may protect individuals for decades or even an entire lifetime. Cross-immunity is acquired protection against one species through contracting an infection from another species. The suppression effect may be mediated by competition for host cells or nutrients, or by heterologous immunity (more details in Section 7.7.6).

7.2. Prospectus

This chapter describes the study area and project phases of the original Garki project. In order to further explore the interactions between *Plasmodium* spp, the Garki data were analysed with two approaches:

1. Cross-sectionally, where the presence of *P. falciparum* in each survey was considered as a risk factor for the presence of the other species in the same survey (Section 7.6.2), and the odds ratios (ORs) between *P. falciparum* and other *Plasmodium* spp were computed. The effects of age, season and locations were assessed. Since most of the papers based on cross-sectional data have explored the associations between species with the same approach, the results of this section could be expected to be comparable with their findings.

2. Longitudinally, where the probabilities of positive slides for *P. malariae* or *P. ovale* were evaluated based on the presence of *P. falciparum* in previous surveys (Section 7.6.3). Furthermore, the effects of *P. falciparum* on the daily acquisition and clearance rates of the other *Plasmodium* spp were explored.

Although the longitudinal approach may be more appropriate to address certain research questions, comparing the results of these two approaches helps explain at least some differences between the results of cross-sectional and longitudinal studies (Section 7.7.3). In addition, the combination of cross-sectional and longitudinal results may help explain in a more comprehensive manner both the short and long-term suppressive effects as well as the acquired immunity between *Plasmodium* spp (Section

7.7.4). Sections 7.6.2 and 7.6.3 present the results of the cross-sectional and longitudinal analyses, respectively; and their differences are discussed in Section 7.7.3.

7.3. Objective

The objective of this part of the thesis is to measure the associations between *P. falciparum-P. malariae* and *P. falciparum-ovale*; assess the effects of repeated infections (i.e., within subject clustering), age, spatial and temporal distribution of individual species on their interactions; and explore the source of associations.

7.4. The original Garki project

The Garki project was one of the largest epidemiological studies on malaria, with data comprised from more than 12,000 people in 23 rounds of treatment. It was conducted in a highly endemic area in northern Nigeria from 1969 to 1976 by co-operation between the World Health Organisation (WHO) and the Nigerian government. This following information is primarily derived from the definitive book by Molineaux and Gramiccia on the Garki project entitled, 'The Garki Project, research on the epidemiology and control of malaria in the Sudan Savanna of west Africa' *[129]*.

7.4.1. Research area

The Garki project research site was situated 500m above sea level and was arid, with one river and no permanent still surface water. The area is considered part of the tropical continental region due to its wide annual and diurnal temperature ranges and restricted rainfall (250-1000 mm per year). The yearly maximum temperature proceeded the rainy season (around 40°C) in March or April, and a secondary peak of high temperature occurred at its end, in October. There were three seasons: dry cool (November-February), dry hot (March-May) and wet (June-October). In the wet season, the research site had temporary collections of water and marshes of various sizes, but these water sources completely dried up in the latter part of the dry season.

The population density of the Garki project area was relatively high. Most residents of the area lived a sedentary life and were farmers; a small group of the population were nomadic herdsmen.

7.4.2. Framework of the original Garki Project

The Garki project had the following objectives:

1. Study the epidemiology of malaria - primarily though measurement of entomological, parasitological and sero-immunological variables and examining their relationships.

2. Measure the effect of specific interventions - an insecticide (propoxur) was used with or without mass drug administration, at two time intervals with different combinations of sulfalene and pyrimethamine.

3. Construct a mathematical model for malarial transmission - this was developed to link entomological and parasitological variables, particularly Vectorial capacity and the prevalence of *P. falciparum.* Also, it could predict the effect of changes in entomology on parasite density, as well as the effect of mass drug administration schemes *[201].*

To achieve its objectives, this project had four successive phases, as follows:

1. Preparatory phase (September 1969 to September 1970). The research protocol was written, the study area was selected, the required forms were developed and the data collection methods were checked.

2. Baseline phase (October 1970 to March 1972; i.e., two dry and one wet seasons). Baseline entomological, parasitological, immunological and meteorological data were collected.

3. Intervention phase (April 1972-October 1973; i.e., two wet and one dry seasons). Epidemiological data were collected in both intervention and control clusters of villages.

4. Post-intervention phase (November 1973-February 1976). Fever cases received active and passive drug administration in some of the villages covered by the mass drug administration programme of the

intervention phase (selective drug administration). Epidemiological data collection was continued in nearly all the villages, most of which had interventions as well as one cluster of untreated control villages.

Regarding the interventions, villages in three concentric areas were treated with one of the following three control strategies (Figure 7-1): insecticide (Area B), insecticide with a low frequency of mass drug administration (Area A2) and insecticide with a high frequency of drug administration (Area A1); each intervention was given in the largest, middle size and smallest areas, respectively *[129]*. Table 7-1 provides information on the number of villages, population within each village, and surface area of each intervention area.

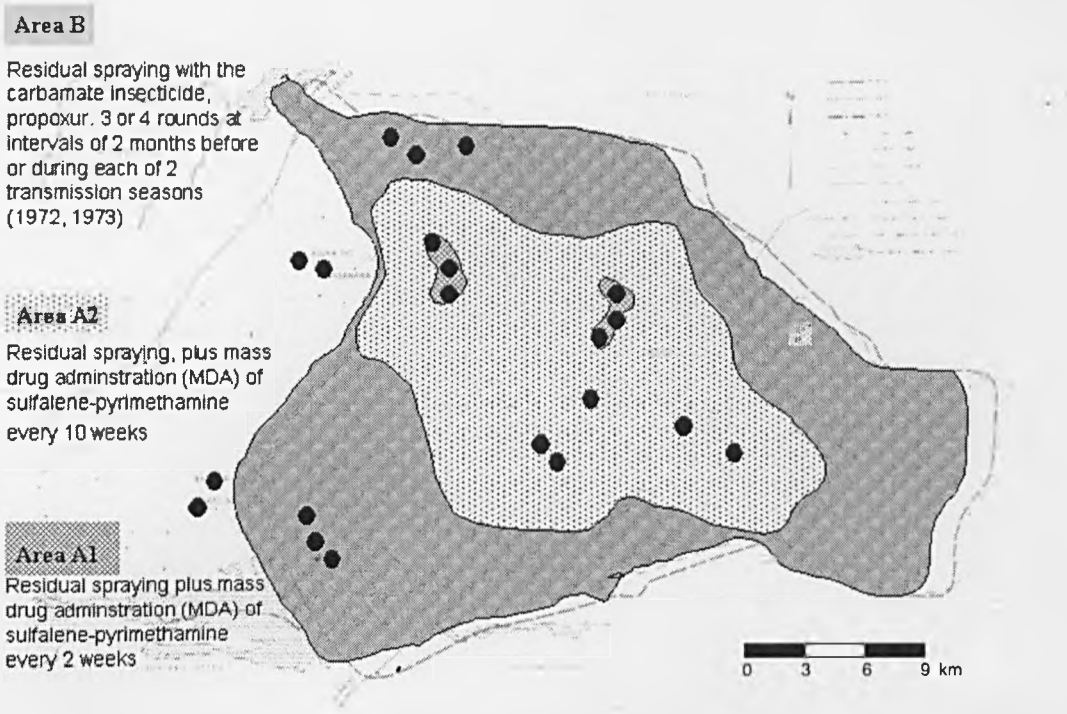

Figure 7-1 Study design of Garki project: the black circles show villages; the four outside the shaded (intervention) area are the control villages without any interventions (the Molineaux and Gramiccia book, page 24)

Table 7-1: Numbers of villages, population and surface in areas treated by the 3 different control strategies (the Molineaux and Gramiccia book, page 28)

Area	Treatment	Number of villages	Population	Surface km²	Period of observation
B	Insecticide[1] alone	104	32,858	550	1970-73
A2	Insecticide + low frequency MDA[2]	54	14,129	350	1970-76
A1	Insecticide + high-frequency MDA[3]	6	1,810	12	1970-73

1 Propoxur
2 Propoxur + MDA every 10 weeks in 1972-3, chloroquine to self-reporting fever cases in 1974-5
3 Propoxur + MDA every 2 weeks wet season and every 10 weeks in dry seasons in 1972-3, chloroquine below 10 years of age every 5 weeks, in wet season in19 74, chloroquine to self-reporting fever cases in19 74-75

7.4.3. Data collection method

Demographic and parasitological data were collected in 23 surveys. These surveys covered the total population of selected village clusters, and were administered every 10 weeks from 1970-73; some intervals between data collection between 1974 and 76 were longer than 10 weeks.

At each survey, a thick blood film was collected and examined for the presence of *Plasmodium* spp in 200 microscopic fields. The numbers of fields that were positive for *P. falciparum* asexual and gametocyte forms; *P. malariae* and *P. ovale* were recorded.

7.5. Data manipulation and analysis

The full data set from the Garki project was received by the Swiss Tropical Institute, Department of Public Health & Epidemiology, in "dbf" format. The file contained 138,197 records of 12,849 subjects in 23 surveys, with each record containing around 50 fields. The main variables of interest for this analysis were the parasitological findings, age, data collection date, and location of each subject in the dataset. Some data related to the non-intervention groups were selected from the full dataset to address the objectives of this chapter (Section 7.6.1).

7.5.1. The cross-sectional analysis

Section 7.6.2 discusses the associations between species in each survey; the ORs were estimated by logistic regression. Since the quadrature assumptions of the accuracy of the numerical integration were met, the random effects model was used to adjust the confidence intervals for the effects of repeated measurements *[202]*. In these analyses, *P. malariae* and *P. ovale* were treated as dependent variables and the main effects and interactions of *P. falciparum,* age, location, season and *P. falciparum* density were assessed. Since this analysis is cross-sectional, the stratum specific ORs (for age and season) of simple cross-tabulation for *P. falciparum* as risk factor for other species can be interpreted the other way around, treating the other species as risk factors for *P. falciparum.*

7.5.2. The longitudinal analysis

Data were manipulated using SQL (Structured Query Language); the queries were written and run in Microsoft Access. In the new format, the parasitological findings of each subject in two consecutive surveys were treated as a record. Hence, some survey data were occurred twice, once as the endpoint of an inter-survey period, and again as the baseline of the following period.

For the purpose of this analysis, only records with both surveys outside the intervention period were included. Nevertheless, the statistical approach was similar to the cross-sectional analysis. In most of the analyses, a single-species infection at the beginning of the initial survey was counted as the risk factor for single or mixed species at the following one. Mixed

infected slides in the initial survey were excluded from the analyses. (Section 7.6.3)

In the next step, for each species, transition frequencies between consecutive surveys were determined, i.e., the numbers N++, N+-, N-+ and N--; where N++ is the number of persons positive at both surveys, N+- the number positive at the first survey and negative at the second one, etc. From these transition frequencies, daily rates of transitions between negative and positive states were derived *[203]* (Appendix 2).

Daily transitional rates were computed in Microsoft Excel. SPSS10 was used for data cleaning and creating new variables, and statistical analyses were done using Stata7.

7.6. Results

The results are presented in Section 7.6.1, from the parasitological findings of the whole dataset to addressing the differing effects of the interventions on the *Plasmodium* spp interactions. Section 7.6.1.1 describes the data from the non-intervention subsets who were eligible for cross-sectional and the longitudinal analysis. The spatial and temporal variations of species are compared in Sections 7.6.1.2 and 7.6.1.3, respectively. Section 7.6.2 illustrates the results of the cross-sectional approach (i.e., the associations between *P. falciparum* with *P. malariae* and *P. ovale* in the same survey) and discusses the effects of repeated observations, age and *P. falciparum* density on species interactions (Sections 7.6.2.5 and 7.6.2.8) and their temporal and spatial variations (Sections 7.6.2.3 and 7.6.2.4). Finally, the

results of longitudinal analysis (i.e., the effect of *P. falciparum* in each survey on the presence of *Plasmodium* spp in the following survey) are given in Section 7.6.3, along with the effects of age and season (Sections 7.6.3.2 and 7.6.3.3). At the end, Section 7.6.3.6 illustrates the effects of *P. falciparum* on the acquisition and clearance rates of other species.

7.6.1. Overview of malaria infection frequency in the complete dataset

The parasitological data of 118,346 out of 138,197 slides were recorded (14.4% of the data were missing). Among the available data, 37% of slides were positive for *P. falciparum*, 10.8% were positive for *P. malariae* and 1.2% were positive for *P. ovale*. The frequency of co-infection with *P. falciparum* and *P. malariae* was 8.5%, while for *P. falciparum* and *P. ovale* it was 0.1%; a total of 435 slides were positive for all species. *P. vivax* did not exist in the Garki area.

In terms of seasonal and spatial deviations, wide ranges of variations were observed; these are further assessed in Sections 7.6.2.3 and 7.6.2.4.

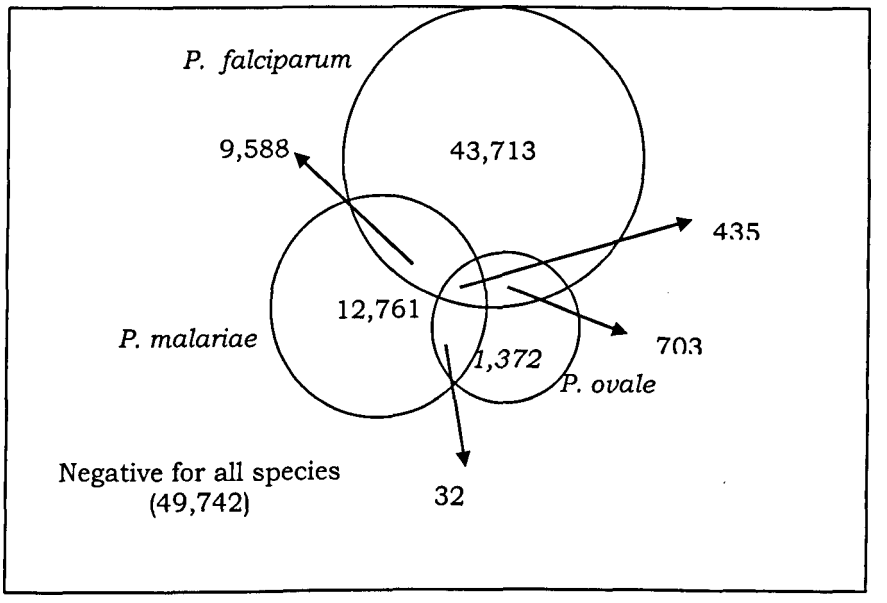

Figure 7-2: Frequencies of single and mixed *Plasmodium* spp in 118,346 assessed slides

Table 7-2 shows the effect of interventions on *Plasmodium* spp prevalences. Frequencies of each species in the different intervention arms were compared with the control group, which contained all records of the non-intervention group and other groups outside the intervention period.

Table 7-2: The effect of interventions on the frequencies of *Plasmodium* spp: the number and percentage of positive slides are classified by intervention and species

Intervention	*P. falciparum*		*P. malariae*		*P. ovale*	
	Number %	Crude OR Adj. OR[2]	Number %	Crude OR Adj. OR[2]	Number %	Crude OR Adj. OR[2]
Non intervention[1]	32,579 49.2	1 1	9,630 14.5	1 1	1,205 1.8	1 1
Insecticide	5,308 35.2	0.56 (0.5-0.6) 0.37 (0.3-0.4)	2,409 16.0	1.1 (1.06-1.2) 1.1 (1.05-1.2)	72 0.5	0.26 (0.2-0.33) 0.22 (0.17-0.28)
Insecticide + high MDA	1,034 7.4	0.08 (0.08-0.09) 0.044 (0.04-0.05)	292 2.1	0.12 (0.11-0.14) 0.09 (0.08-0.1)	28 0.2	0.11 (0.07-0.16) 0.09 (0.06-0.13)
Insecticide+ low MDA	665 7.5	0.08 (0.08-0.09) 0.037 (0.03-0.41)	106 1.2	0.07 (0.06-0.08) 0.05 (0.04-0.06)	8 0.1	0.05 (0.02-0.09) 0.04 (0.02-0.08)
Selective DA[3]	4,127 29.4	0.43 (0.41-0.45) 0.31 (0.3-0.33)	324 2.3	0.14 (0.12-0.15) 0.1 (0.09-0.11)	59 0.4	0.22 (0.17-0.29) 0.2 (0.16-0.26)
Total	43,713 36.9	(rho = 0.45)	12,761 10.8	(rho = 0.44)	1,372 1.2	(rho = 0.35)

1. Non-intervention group contained all records of non-intervention group and the records of other groups outside the intervention period
2 Adjustment of repeated observations; person were treated as cluster, and rho showed inter-cluster correlation
3 The drug was administrated to fever cases in post-intervention phase in MDA arms (Section 7.4.2)

From Table 7.2, it is evident that insecticide alone had no preventive effect on the *P. malariae* (adjusted OR=1.1), however it decreased the risk of *P. falciparum* and *P. ovale* dramatically (adjusted OR=0.37 and 0.22, respectively). On the other hand, the combination of insecticide and mass drug administration had a similar effect, more or less, on every species.

Because each of the interventions had a different impact on the frequencies of *Plasmodium* spp, it is also plausible that they had unbalanced effects on

the patterns of those interactions, particularly if the effects of drug administration coverage and resistance are also taken into account. Therefore, it was decided that the interactions between *Plasmodium* spp would only be assessed in the non-intervention sets (i.e., all records of the control group) and pre and post intervention phases of the other groups (Section 7.6.1.1).

7.6.1.1. Description of the non-intervention data

In Table 7-3, the numbers of eligible records for cross-sectional and longitudinal analysis are shown in different project phases and intervention arms. It contains all records in the pre-intervention phase, some records from the post-intervention phase and all records of the non-intervention arm. Since the first post-intervention survey was done around 2-3 months after the end of interventions, it could be expected that most effects of the interventions were washed out.

In the longitudinal analysis, only those records were included in which both former and latter surveys were not in the intervention phase. Hence, the numbers of eligible records were not exactly the same in these two analyses (71,270 records in cross-sectional and 68,894 ones in longitudinal analysis).

Table 7-3: Number of selected records according to the project phases

Follow-up group		Phase of data collection			Total
		Pre-intervention	Main intervention	Post-intervention	
Control (no intervention)	Cross-sectional	14,120	13,915	8,476	36,511
	Longitudinal	12,969	12,363	7,071	32,403
Insecticide alone	Cross-sectional	13,698	0	0	13,698
	Longitudinal	13,170	0	0	13,170
Insecticide + high-frequency MDA[1]	Cross-sectional	11,441	0	1,080	12,521
	Longitudinal	10,990	0	4,188	15,178
Insecticide + low-frequency MDA[1]	Cross-sectional	8,540	0	0	8,540
	Longitudinal	8,143	0	0	8,143
Total	Cross-sectional	47,799	13,915	9,556	71,270
	Longitudinal	45,272	12,363	11,259	68,894

1 Mass drug administration

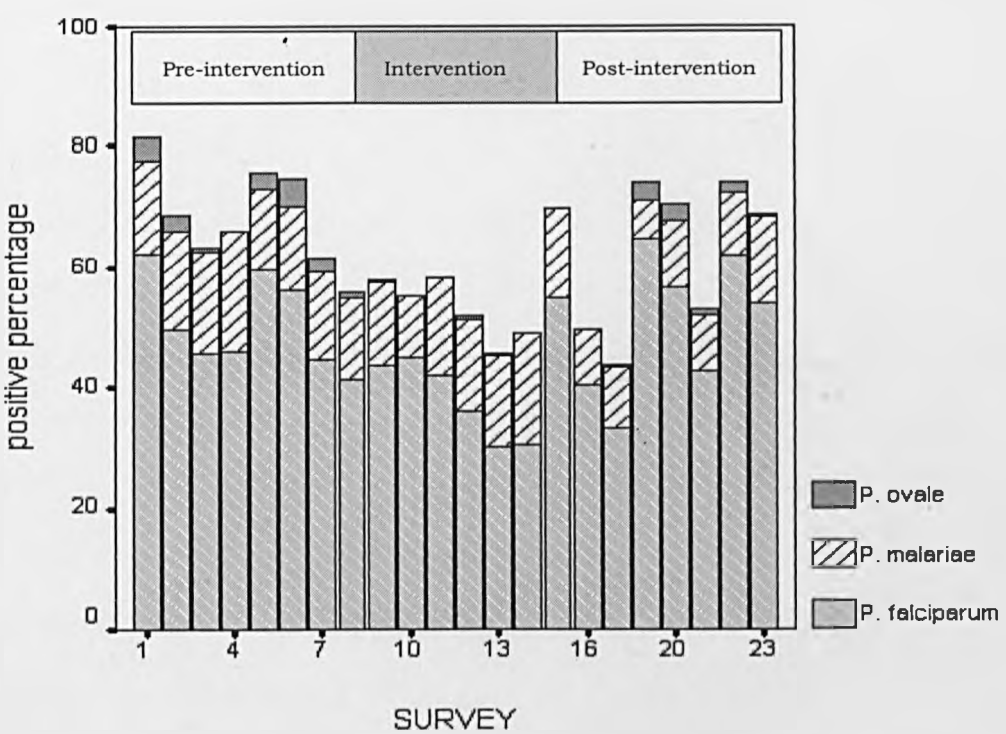

Figure 7-3: Frequencies of *Plasmodium* spp in every survey, mixed infected slides were counted more than once.

Figure 7.3 illustrates that the relative frequencies of species are more or less comparable in surveys and more than 80% of positive slides were positive for *P. falciparum*.

7.6.1.2. Between village variations of the infections

The largest village in the study had a population of 1,130 people while the smallest one had a population of 92 people. On average, villagers were surveyed 6.1 times (SD=3.3). Comparing villages, the smallest average number of surveys per person was 2.6; this number was 13.2 times as great for a village with the longest follow up (this village belonged to the non-intervention arm and all data of its three phases were included in analyses).

Among collected slides, 50% were positive for *P. falciparum* (SPR), 15% for *P. malariae* and 3% for *P. ovale* (Table 7-4). However, the correlation structure of data was taken into account in multilevel models in order to estimate within- and between-village variations and adjust for the clustering effect of individuals, surveys and villages. Using MLwiN 1.1, a 3-level model was considered for each species; survey number, person ID and village code were the first, second and third levels, respectively. The models fitted the logit of species based on the random effect of the first level, and fixed and random effect of the second and third levels (Appendix 3). Hence, the estimated fixed constant shows the logit of SPR, and the random effects of the second and third levels show within- and between-village variations. To optimise goodness of fit, the Restricted Iterative Generalised Least Squares (RIGLS) method was applied where the convergence of each parameter was assessed individually *[204]*.

Table 7-4 shows the results of these multilevel models. The adjusted SPR for *P. falciparum*, *P. malariae* and *P. ovale* were 53.6%, 12.4% and 2.2%, respectively. The results of heterogeneity tests show highly significant differences in the within- and between-village SPRs; these variations were more or less comparable among species.

Table 7-4: The estimation of species' SPRs, and their within- and between-village variations, applying multilevel analysis

Population of villages	
Mean (SD)	469.5 (253)
Minimum- maximum	92-1130
The number of survey rounds[2]	
Mean (SD)	6.1 (3.3)
Minimum- maximum	2.55-13.2
Crude SPR[1] (%)	
P. falciparum	50
P. malariae	15
P. ovale	3
Adjusted SPR for hierarchical structure of data ($e^{\beta 1}$) (%)	
P. falciparum	53.6
P. malariae	12.4
P. ovale	2.2
Between village variations of SPR (σ^2_{v1})	
(95% CI of village SPR)	
P. falciparum (χ^2, p-value[3])	51-56 (7.43,0.006)
P. malariae (χ^2, p-value)	12-14 (5.55, 0.018)
P. ovale (χ^2, p-value)	1-3 (8.77, 0.003)
Within village variations of SPR (σ^2_{u1})	
(95% CI of individual SPR)	
P. falciparum (χ^2, p-value)	2-98 (2004, <0.0001)
P. malariae (χ^2, p-value)	0.7-74 (1031, <0.0001)
P. ovale (χ^2, p-value)	0.4-10 (51.49, <0.0001)

1: Smear Positive Rate

2: Of the 22 within-village averages of rounds per person

3: Heterogeneity test based on quasilikelihood method, the results of χ^2 with df=1

(Appendix 3 for definitions of σ^2_{u1}, σ^2_{v1}, and $e^{\beta 1}$)

Figure 7.4 shows that there was a wide heterogeneity in the risks of having of positive slide in consecutive surveys, not only between inhabitants of different villages, but also within villages.

The very wide range of within-village SPRs could be due to the heterogeneity in exposure risks; however, at least some part of this wide range might be explained by acquired immunity. Young people with less immunity were more susceptible to repeated infection and/or had a lower parasite clearance rate. Partial acquired immunity in the older age group might protect these individuals from having repeated positive slides *[138]*.

In his seminal book on the Garki project, Molineaux and Gramiccia did not explore within- and between-village variations of infection risks and therefore, the above findings were original. The heterogeneity of repeated infections between villages are discussed in more details in Section 7.6.2.4.

7.6.1.3. Temporal variations of the infections

Molineaux and Gramiccia discussed, in general, seasonal variations of *Plasmodium* spp *[138]*. However, this section explores this issue more widely and explains the results in relation to the results of the previous sections in order to explore the possible effects of season on the interaction between species.

The monthly variation in SPR is illustrated in Figure 7-4. The maximum and minimum SPR of *P. falciparum* were 64% and 39%, respectively. The corresponding figures for *P. malariae* and *P. ovale* were 21%, 11% and 5%,

0%, respectively; the annual *P. falciparum* and *P. ovale* variations were more or less comparable. The SPRs of both of these species peaked at the end of the wet season (September to November) when the mosquito prevalence was highest, and the lowest SPRs were detected in the dry-hot season (April to June). However, the temporal variation of *P. malariae* had exactly the reverse pattern, i.e., its highest frequency was detected in the dry-hot season, in which other species and the mosquito population were at their minimum densities. This difference in SPR peaks is very important, because a direct association between the population of vectors and the frequency of all *Plasmodium* spp would be expected.

Molineaux and Gramiccia showed that the *P. malariae* infant conversion rate, the rate at which infants became positive for the first time, had a positive association with vector density and had its highest magnitude in wet season. The discrepancy between seasonal variations of *Plasmodium* spp were not detected in children younger than 5 years old *[138]*. Therefore, it is unlikely that the different temporal patterns of species were due to either genetically determined characteristics of the parasites or the effect of environmental factors on the extrinsic cycles of parasites.

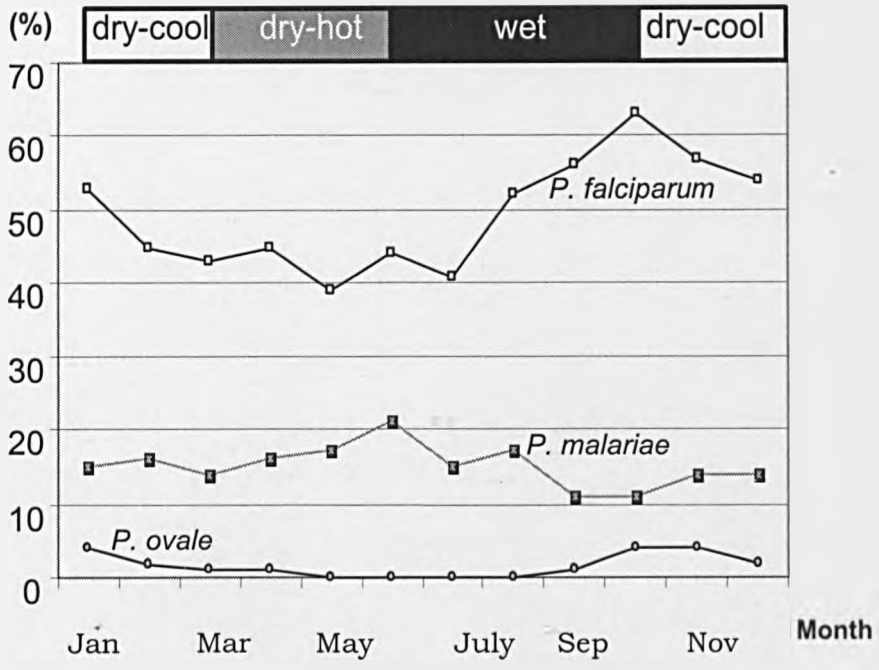

Figure 7-4: Annual variation of *Plasmodium* spp prevalence, based on 6 years of slide data

Alternatively, the lack of concurrent increases in the prevalences of vectors and *P. malariae* could be explained by the suppressive effect of *P. falciparum* on the patent form of *P. malariae* in the wet season. Suppression of one species of *Plasmodium* by another is known from clinical observations *[205]*. Looareesuwan et al. (1987) found that among patients treated for acute *P. falciparum, P. vivax* developed in one third of patients within one month of receiving a regimen containing quinine or quinidine or after two months of receiving mefloquine treatment; this was much higher than the risk of infection in normally found in Thailand, and much higher than the normal risk of mixed infection *[135]*. Based on this finding, he concluded that *P. falciparum* suppressed *P. vivax*, and furthermore that

after treatment of *P. falciparum*, *P. vivax* became patent. The prevalence of mixed infections was higher based on molecular diagnostic techniques than the microscopic diagnosis commonly used in western Africa, and again showed that a considerable portion of subjects had both infections but by microscopy they showed only *P. falciparum [132].*

Based on the Garki data, suppression of *P. malariae* by *P. falciparum* is also suggested by the timing of observed events. A rapid increase in vector density was apparent after the onset of the wet season and was rapidly followed by a marked increase in the prevalence of *P. falciparum.* This coincided with a marked decease in the prevalence of *P. malariae*, which was observed repeatedly in different villages over several years (Section 7.7.6).

Table 7-5 shows that compared to *P. malariae*, *P. falciparum* has a shorter pre-patent period, which might cause the frequency of *P. falciparum* to elevate faster by increasing mosquito density in the wet season while having a suppressive force on other species, particularly *P. malariae.* In addition, as *P. malariae* has lower blood density, it could be suppressed by *P. falciparum* to undetectable levels faster than *P. ovale.* To more fully understand the force of suppressive effects between species, cross-immunity between species must also be taken into account (Section 7.7.5).

Table 7-5: Some characteristics of *Plasmodium* spp infections in humans (Gilles (1993)) *[1]*

	Species			
	P. falciparum	*P. malariae*	*P. ovale*	*P. vivax*
Pre-erythrocytic stage (days)	5.5-7	14-16	9	6-8
Pre-patent period from inoculation (days)	9-10	15-16	10-14	11-13
Incubation period (days)	12 (9-14)	28 (18-40) or longer	17 (16-18) or longer	15 (12-17) or up to 6-12 months
Erythrocytic cycle (hours)	48	72	50	48
Parasitaemia per mm^3				
Average	20,000-500,000	6,000	9,000	20,000
Maximum	2,000,000	20,000	30,000	50,000
Primary attack	Severe	Mild	Mild	Mild to severe
Febrile paroxysm (hours)	16-36 or longer	8-10	8-12	8-12
Relapses	-	-	++	++
Period of recurrence	Short	Very long	Variable	Variable
Duration of untreated infection (years)	1-2	3-50	1.5-5	1.5-5

7.6.2. Cross-sectional analysis

Although the interactions between species were analysed cross-sectionally in the Garki project and some analyses have been published, most of the results of this section, particularly discussions and relationships between the findings, are original and are explained in detail in the following sections.

This section presents the frequency of dually infected slides and explains effect modification of age, season and spatial distribution by species on the

mixed infections (Section 7.6.2.4). Having reviewed the crude odds ratios (ORs) between *P. falciparum* (as the independent variable) and other species, the effect of repeated observations is discussed in Section 7.6.2.2. Then, the temporal and spatial variations of dually infected slides and the effects of age and infection density are quantified in Sections 7.6.2.6 to 7.6.2.8.

7.6.2.1. Crude associations

Of the 43,713 *P. falciparum* positive slides, 23.8% and 3.2% were positive for *P. malariae* and *P. ovale,* respectively. Only 5.7% and 0.6% of the 31,235 negative *P. falciparum* slides were positive for *P. malariae* and *P. ovale,* respectively (*P. malariae*: χ^2=407, p<0.0001 and OR=5.2; and *P. ovale* χ^2=556, p<0.0001 and OR=5.7). Based on these results, *P. falciparum* increased the risk of other infections more than 5 times (Table 7-6). These crude ORs were calculated based on repeated observations, therefore, before drawing any conclusion, they should have traditionally be adjusted. However, their high magnitudes showed that the risk of *P. falciparum* and other *Plasmodium* spp were strongly correlated.

7.6.2.2. Repeated measurement effect

Using a logistic regression model adjusted for repeated observation by random effects, the adjusted ORs between species were computed (Section 7.5). The adjusted OR of *P. falciparum* as a risk factor for *P. malariae*

decreased from 5.2 (crude OR) to 3.6 (95% CI: 3.4-3.9). The corresponding figures between *P. falciparum* and *P. ovale* decreased from 5.37 (crude) to 5.1 (95%CI: 4.3-6). In addition, the intra-cluster correlation coefficients (rho) in simple models without any predictors (in each logistic regression model, only one species was entered) for *P. falciparum, P. malariae* and *P. ovale* were 0.55 (CI: 0.5-0.53) 0.46 (CI: 0.44-0.47) and 0.39 (CI: 0.34-0.43), respectively (Section 7.5 for explanation).

The adjustment decreased the association between *P. falciparum* and *P. malariae* dramatically; although it did not have much effect on the association between *P. falciparum* and *P. ovale.* In addition, it showed that intra-person correlation was higher in *P. malariae* than *P. ovale;* intra-person correlations for both were smaller than that for *P. falciparum.*

Table 7-6: The association of *P. falciparum* (as risk factor) with other species, classified by age and season; i.e., the ORs are stratum specific

Model		*P. malariae* OR (95% CI)	*P. ovale* OR (95% CI)
Crude	All subjects	5.2 (4.92-5.49)	5.37 (4.6-6.28)
	Age (year)		
	<1	16.62 (9.42-29.32)	6.25 (2.63-14.82)
	1-9	3.35 (2.95-3.79)	2.32 (1.70-3.16)
	>=10	2.69 (2.50-2.89)	3.97 (3.24-4.85)
	Season		
	Dry and cool	5.59 (5.14-6.08)	6.05(4.97-7.35)
	Dry and hot	6.85(5.90-7.95)	3.92(2.17-7.08)
	Wet	4.39(4.02-4.81)	3.73(2.76-5.03)
Adjusted for repeated observations*	All subjects	Rho**=0.34 3.64 (3.4-3.9)	Rho**=0.25 5.1 (4.33-6.0)
	Age (year)		
	<1	6.25 (2.63-14.82)	6.26 (2.64-14.83)
	1-9	2.32 (1.70-3.16)	2.19 (1.59-3.03)
	>=10	3.97 (3.24-4.85)	3.95 (3.23-4.84)
	Season		
	Dry and cool	4.02(3.7-4.35)	5.53(4.6-6.68)
	Dry and hot	6.32(5.48-7.29)	3.94(2.18-7.12)
	Wet	3.58(3.3-3.9)	3.76(2.78-5.07)

* Logistic regression models with random effect were applied. None of them were multivariate model. In each step, only one independent variable was entered and stratum specific ORs were computed

** A measure of the intra-person clustering effect ranges between 0 and 1

The highest heterogeneity in *P. falciparum* risk means that the variation in susceptibility to *P. falciparum* was wider compared to the other species. In other words, a group of people was contracting *P. falciparum* repeatedly, while another group of people had greater protection against it. The range of susceptibilities of people to *P. falciparum* was greater than it was for people with either *P. ovale* or *P. malariae*. This is compatible with the finding that the within- and between-village SPR variations in *P. falciparum* were greater than the variations in other species (Section 7.6.1.2).

Nevertheless, the greater heterogeneity in the risk of repeated *P. falciparum* infection could not be explained simply by the differences in the biology of

Plasmodium spp; because *P. falciparum* does not generate hypnozoites, it does not produce relapses (Table 7-5). In short, there are no convincing explanations for any possible differences in the transmission risks heterogeneities among *Plasmodium* spp, except acquired immunity which could explain most of the findings (Section 7.7.4).

7.6.2.3. Temporal variations of odd ratios

The OR between *P. falciparum* (as a risk factor) and *P. malariae* adjusted for season and repeated observations was 2.97 for the whole year; season specific ORs in the dry-cool, dry-hot and wet seasons were 4.02, 6.32 and 3.58, respectively. The corresponding OR between *P. falciparum* and *P. ovale,* again adjusted for season and repeated observations, was 4.9 for the whole year; season specific ORs were 5.53, 3.94 and 3.76 for the dry-cool, dry-hot and wet seasons respectively.

Figure 7-5 shows wide temporal variations between the ORs of *P. falciparum* and other species. The results of heterogeneity tests by the Mantel-Haenszel method showed significant differences between the monthly and seasonal ORs (all heterogeneity test p-values were less than 0.0001) However, within each season, the monthly ORs were more or less comparable. The peak in OR between *P. falciparum* and *P. ovale* was at the end of the dry-cool season with a very wide confidence interval (6.1-41).

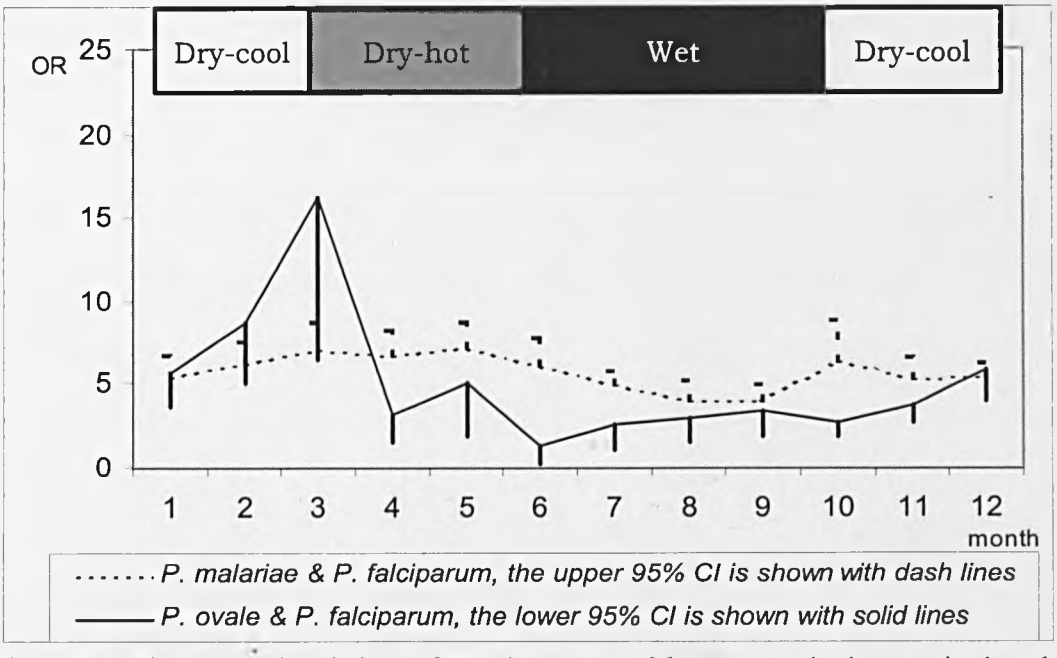

Figure 7-5: The temporal variations of ORs between *P. falciparum* and other species based on 6 years of data

The highest ORs between *P. falciparum* and *P. ovale* were detected in the dry-cool season, just after the peaks of both species, and were compatible with the corresponding figure in the longitudinal analysis (Section 7.6.3.3). The highest ORs between *P. falciparum* and *P. malariae* were detected at both the end of the dry-cool and beginning of the dry-hot seasons; these peaks were after the *P. falciparum* peak and before that of *P. malariae* (Figure 7-4). This means that because *P. falciparum* decreased in the dry-cool season, the OR between these two species increased, and then *P. malariae* increased in the dry-hot season (see following paragraphs and Section 7.7.6 for its explanation).

Molineaux and Gramiccia found that higher infant conversion rates were observed for all species in the wet season compared to the other seasons *[138]*. Conversely, in those aged more than 5 years, *P. malariae* infection was less common in wet season. This temporal sequence is compatible with

the hypothesis that *P. falciparum* exerts a suppressive effect on *P. malariae* (Section 7.7.6). This suppression hypothesis implies that as the density of *P. falciparum* infection declined in the dry-cool season, sub-patent *P. malariae* became patent in mixed-infection slides; however, later on, a higher number of single *P. malariae* slides were detected due to greater *P. falciparum* clearance and low transmission. However, more evidence is needed to test this hypothesis; longitudinal analysis further examined this issue (Section 7.6.3.3).

The mean age of people infected with only *P. falciparum* was 17.2 years, while the corresponding means for *P. ovale* and *P. malariae* were 26 and 29 years (greater than the average age in the whole population, which was 24.6 years). The mean age of people infected with both *P. falciparum* and *P. ovale* was 11 years, and that of people infected with *P. falciparum* and *P. malariae* was 10.4 years. These figures show that single *P. ovale* and *P. malariae* infections were most often patent in elderly people; however, mixed infections were more common in younger age groups. This age pattern was more or less constant in all seasons.

The above finding about the age distributions of infected people with single and mixed species could also be explained by the effect of acquired immunity. Younger people, who were less exposed to infections might have had less protection against all species, as well as may have been more susceptible to acquire different species. Specifically, the evidence that *P. falciparum* is associated with a lower mean age of infection is consistent with theories that it produces acquired immunity.

The above explanation is also compatible with the higher observed heterogeneity risk of *P. falciparum* infection compared to the other species (Section 7.6.2.5). A higher frequency of infections in the early years of life and lower infection risk in adulthood, probably due to acquired immunity, increased the differences of *P. falciparum* infection risks between young and adult groups. This might have increased the observed heterogeneity risk.

7.6.2.4. Spatial variations of odd ratios

Among 22 villages, the minimum and maximum of ORs between *P. falciparum* and *P. malariae* were 3.91 and 8.87, respectively (heterogeneity test: $\chi^2=65$, $p<0.0001$). The corresponding figures between *P. falciparum* and *P. ovale* were 1.29 and 18.53, respectively (heterogeneity test: $\chi^2=33.4$, $p=0.04$). These results show that the ORs between *P. falciparum* and other species were not uniform among villages.

To explore the possible sources of these heterogeneities, an ecological analysis was performed. Also, the Pearson correlation coefficients between the ORs and the demographic and malarialogical factors of villages (population, duration of follow up period, SPRs, and their within village variations) were computed. Although none of these correlation coefficients were significant (all p-values were greater than 0.4), small negative associations were detected between the ORs and SPRs, suggesting that greater ORs were seen in villages with lower averages of smear positive rates. However, after adjustment for these factors, the heterogeneity still existed.

This finding was compatible with the observed negative associations between the frequencies of species and the ORs between species in meta-analysis (Section 5.5.6).

7.6.2.5. The age effect

Table 7-7 presented the risk of infections classified by age (less than 1, and 1-9 years and more than 10 years old). Having put the infection risk of more than 10 years olds as the baseline, the 1-9 year old group had the highest ORs, particularly with *P. falciparum* (11.68 in *P. falciparum* versus 5.9 and 4.2 in *P. malariae* and *P. ovale* respectively). It means that the relative risk of *P. falciparum* in children less than 10 years old versus the older age group was more prominent compared to other species risks.

There is no evidence to support lower exposure risk to infections in the older age group *[18]*. Therefore, acquired immunity against species is the only reasonable explanation for higher infection risks in children less than 10 years old.

The higher *P. falciparum* OR could be explained by stronger natural acquired immunity. Having experienced more *P. falciparum* infections during the first decade of life, people were less susceptible afterwards. This idea is also supported by a lower mean age of apparent infected subjects with *P. falciparum* compared to the other species (Section 7.6.2.3). However, it should be noted that the booster effect of repeated inoculations of *P. falciparum* in older age might have a major role in keeping up the alertness of the immune system even in non-patent forms.

In addition, the suppressive effect of *P. falciparum* on the other species might also explain some of the differences between the ORs. Given this fact, *P. falciparum* as the dominant species might not allow others to be patent in the 1-9 year-old age group. However, after age 10 years, as *P. falciparum* incidence declined, other species would find more chance to be patent, therefore the relative risk of infection between ages 1-9 years and more than 10 years was less prominent in *P. malariae* and *P. ovale*.

7.6.2.6. Effect modification by age

Regarding the interactions between *P. falciparum* and other species, there was a wide heterogeneity between age stratum specific ORs. The ORs between *P. falciparum* and *P. malariae* were 25.4, 3.03, and 2.32 in less than one year, one to nine years and more than nine years, respectively. The corresponding figures for *P. falciparum* and *P. ovale* were 6.26, 2.19 and 3.95, respectively. As the above ORs show, there were stronger positive associations between *P. falciparum* and other species in children aged less than one year old.

Based on the heterogeneity hypothesis (Chapter 6), the differences between the age-specific ORs could be explained as follows. Since infants had less history of infections and acquired less immunity, the risks of infections were highly correlated to the exposure risk; i.e., heterogeneity in exposure risk increased the overall OR. However, older people might get more immunity due to previous infections; therefore, the risks of infections were less strongly linked to the exposure. In other words, highly exposed people might be less susceptible to infections due to higher acquired immunity,

and vice versa. Hence, lower ORs might be expected as the heterogeneity in the infection risks were less than the heterogeneity in the exposure risks.

7.6.2.7. Acquired maternal immunity

Table 7-7 compared the age-specific infection risks; the baseline was the risk in more than ten year old group. Lower observed risk of infections in less than one year olds and the 1-9 age group (ORs were 2.1 and 11.7 for *P. falciparum*, 1.3 and 5.9 for *P. malariae*, and 2 and 4.2 for *P. ovale*) might be due to either less exposure to vectors, and/or maternal acquired immunity, which partially protected babies in the first years of their lives. Kitua et al. (1996) *[206]* and Sehgal et al. (1989) *[207]* showed that maternal immunity might protect babies against severe malaria attacks for a few months if mother got infection or disease during pregnancy, particularly in the second part of pregnancy.

To explore this issue more deeply, the ORs of infections in the first 4, 4-7 and 8-12 months of life were computed based on their risks in children aged ten or more years old. The ORs for babies aged less than four months were 0.75, 0.56 and 1 for *P. falciparum*, *P. malariae* and *P. ovale*, respectively; these were much lower than the ORs for the rest of first year of life (Table 7-7). This was compatible with other findings that four months seems to be the duration of effective maternal acquired immunity *[206,207]*.

Infection risks in infants, particularly those aged less than four months, seemed to be even lower than those for the oldest age group to which their mothers belonged. Having assumed effective maternal acquired immunity, one would not expect to detect stronger protection among infants than that

conferred upon their own mothers. Therefore, this lower infection risk in babies might be due to not only the obtained immunity from their mothers, but also a lower chance of effective contact with vectors as a result of higher protection by their families or less attraction between mosquitoes and babies (lower exposed body surface or different temperature) *[208-210]*.

Table 7-7: The risk of infection with *Plasmodium* spp classified by age; the infection risk in the last age group (=10 years) is the baseline

Age group	<4 months Number (%)	4-7 months Number (%)	8-12 months Number (%)	1-9.9 year Number (%)	=10 year Number (%)	Total
P. falciparum						
Negative	484 (70.7)	356 (42)	272 (31.6)	2,220 (13.5)	27,902 (64.6)	31,234
Positive	201 (29.3)	492 (58)	558 (68.4)	14,224 (86.5)	15,299 (35.4)	30,804
OR (95% CI)	0.75 (0.64-0.9)	2.52 (2.2-2.9)	3.9 (3.41-4.56)	11.68 (11.13-12.27)	1 -	χ^{2*}=1270 p-value <0.001
	OR for the whole first year: 2.1 (1.8-2.4)					
P. malariae						
Negative	654 (95.9)	763 (90)	738 (85.8)	10,957 (66.6)	39,826 (92.2)	52,938
Positive	31 (4.5)	85 (10)	122 (14.2)	5,486 (33.4)	3,375 (7.8)	9,099
OR (95% CI)	0.56 (0.39-0.8)	1.31 (1.04-1.65)	1.95 (1.6-2.37)	5.9 (5.63-6.2)	1 -	χ^{2*}=6285 p-value <0.001
	OR for the whole first year: 1.3 (1.1-1.5)					
P. ovale						
Negative	678 (99)	826 (97.4)	841 (97.8)	15,761 (95.8)	42,761(99)	60,867
Positive	7 (1)	22 (2.6)	19 (2.2)	682 (4.2)	440(1)	1,170
OR (95% CI)	1 0.47-2.12	2.59 (1.68-4)	2.2 (1.38-3.49)	4.2 (3.72-4.75)	1 -	χ^{2*}=630 p-value <0.001
	OR for the whole first year: 4.2 (3.6-5.0)					

*χ^2 tests checks the associations between age group and the frequencies of species

7.6.2.8. The effect of *P. falciparum* density

Based on the number of positive *P. falciparum* fields, out of 200 examined, slides were categorised into three groups, slides with 0, 1-50 and more than

50 positive fields. Then, the risk of infection with *P. malariae* and *P. ovale* in the second and third density groups were compared with the negative (0) group. The ORs in the 1-50 and more than 50 positive fields were 4.05 and 8.66 for *P. malariae*; 4.05 and 8.73 for *P. ovale*. Hence, higher risks of infections were detected in greater density of *P. falciparum* infection.

Based on the above results, a positive association between *P. falciparum* density with the risks of other species were detected. However, even though negative associations might be expected because of observational bias (missing other species in slides infected with a high density of *P. falciparum*), these results show positive relationships. It may seem that this finding is not compatible with the suppression hypothesis. However, Section 7.7.5 discusses how this finding contradicts the explanation that competition for host cells or nutrients is responsible and supports the possible role of immunity in suppression effect.

The observed positive associations might be explained as follows. Given short-term cross-immunity between species, in individuals infected with *P. falciparum*, one may expect to see more *P. malariae* and *P. ovale* infections. Also, the partial immunity against *P. falciparum* usually reduces its blood density *[92];* the Garki project results showed a negative association between IgM level and *P. falciparum* density in all age groups *[211].* Therefore, highly infected slide might show lower immunity against *P. falciparum*, and greater infection risk for the other species.

Nevertheless, as the density of *P. falciparum* and the infection risks decreased by age, the density effect might be confounded by age. Older age groups experienced more infections with all species, therefore, obtained

stronger immunity against every species. Therefore, lower density of *P. falciparum* with less risks of infection with *P. malariae* and *P. ovale* could be as a result of acquired immunity.

7.6.3. Longitudinal analysis

In this part of the analysis, the presence of each species in the former survey was considered as a risk factor for the presence of *Plasmodium* spp in the latter survey, and their crude and adjusted (for the effect of repeated measurements) ORs were computed. It presents the effect of species-specific acquired immunity (Section 7.6.3.4) and their cross-immunities (Sections 7.6.3.1 and 7.6.3.5). Also, the effect modifications of age and season are assessed in Sections 7.6.3.2 and 7.6.3.3, respectively. In the last section (7.6.3.6) the conversions rates (acquisition and clearance) of *P. malariae* and *P. ovale* are computed and the effects of *P. falciparum* on these rates are assessed.

7.6.3.1. *P. falciparum* as a risk factor for other species

Out of 68,894 slides, 54,730 (79.4%) had parasitological data in both former and latter surveys (Table 7-3). Table 7-8 shows the frequency of positive slides based on the results of their baselines. All positive

associations were statistically significant, and show that the infection risks were higher in those who had positive slides in the former survey. However, the final discussion should be based on the adjusted ORs for the effect of repeated observations (Table 7-9).

Table 7-8: The frequencies of infections at the latter survey in relation to the infection status in the former survey

	Latter survey					
	P. falciparum		*P. malariae*		*P. ovale*	
Former survey	Negative	Positive	Negative	Positive	Negative	Positive
P. falciparum						
Negative	20,201	7,934	25,080	1,522	27,802	176
(%)	(71.8)	(28.2)	(94.3)	(5.7)	(99.4)	(0.6)
Positive	7,933	18,662	17,223	3,120	25,147	608
(%)	(29.8)	(71.2)	(84.7)	(15.3)	(97.6)	(2.4)
OR(95% CI)	6(5.7-6.3)		3.0 (2.8-3.2)		3.8 (3.2-4.5)	
P. malariae						
Negative	19,309	7,293	42,303	4,642	45,723	555
(%)	(72.6)	(27.4)	(90.1)	(9.9)	(98.8)	(1.2)
Positive	892	641	4,498	3,287	72,26	229
(%)	(58.2)	(41.8)	(57.8)	(42.2)	(96.9)	(3.1)
OR(95% CI)	1.9 (1.7-2.1)		6.7 (6.3-7.1)		2.6 (2.2-3.1)	
P. ovale						
Negative	20,115	7,863	41,807	447	52,949	784
(%)	(71.9)	(28.1)	(90.3)	(9.7)	(98.5)	(1.5)
Positive	86	71	496	171	887	110
(%)	(54.8)	(45.2)	(74.4)	(25.6)	(89)	(11)
OR(95% CI)	2.1 (1.5-2.9)		3.2 (2.7-3.8)		8.4 (6.8-10.3)	

Records with positive results for the species of interest in the former survey were excluded in this analysis. For example, the association between *P. falciparum* in the former survey and *P. malariae* in the following survey was assessed only in those who were negative for *P. malariae* in the former survey. Each slide (except the first & last one for each person) went in to two records; once as the initial result & once as the subsequent one (Section 7.5.2)

The crude ORs between *P. falciparum* as a risk factor for *P. malariae* and *P. ovale* were 3 (95% CI: 2.8-3.2) and 3.8 (95% CI: 3.2-4.5) respectively. The corresponding adjusted ORs were 2.7 (95% CI: 2.5-2.9) and 3.6 (95% CI: 3-4.4). The ORs between *P. falciparum* and other species in cross-sectional

analysis were greater than the observed ORs in longitudinal analysis. Section 7.7.3 discusses the possible explanation for these differences.

Table 7-9: The ORs between *Plasmodium* infections in two consecutive surveys

Detected species in the former survey	Detected species in the latter survey	Crude OR (95% CI)	Adjusted OR[2] (95% CI) rho
P. falciparum[1]	*P. falciparum*	6.0(5.7-6.3)	1.9 (1.9-2) 0.73
	P. malariae	3.0 (2.8-3.2)	2.68 (2.5-2.9) 0.44
	P. ovale	3.8 (3.2-4.5)	3.6 (3-4.4) 0.34
P. malariae[1]	*P. falciparum*	1.9 (1.7-2.1)	1.7 (1.5-2.0) 0.22
	P. malariae	6.7 (6.3-7.1)	2.7 (2.5-2.9) 0.33
	P. ovale	2.6 (2.2-3.1)	2.6 (2.2-3.0) 0.03
P. ovale[1]	*P. falciparum*	2.1 (1.5-2.9)	1.9 (1.3-2.8) 0.22
	P. malariae	3.2 (2.7-3.9)	2.8 (2.3-3.4) 0.29
	P. ovale	8.4 (6.8-10.4)	5.3 (3.9-7.2) 0.17
P. falciparum+ *P. malariae*	*P. falciparum*	7.3 (6.8-7.9)	2.4 (2.2-2.6) 0.5
	P. malariae	7.9 (7.5-8.4)	3.1 (2.8-3.3) 0.32
P. falciparum+ *P. ovale*	*P. falciparum*	6.4 (5.3-7.7)	2.4 (1.9-3.1) 0.53
	P. ovale	9.8 (7.9-12.1)	6.1 (4.4-8.4) 0.17
Any *Plasmodium* spp	*P. falciparum*	5.8 (5.6-6)	1.9 (1.8-2.0) 0.73
	P. malariae	4.9 (4.6-5.2)	2.9 (2.7-3.1) 0.63
	P. ovale	4.5 (3.8-5.3)	4.2 (3.5-5.0) 0.53
	Any *Plasmodium* spp	5.7 (5.5-5.9)	1.8 (1.7-1.9) 0.72

1: positive for the given species and negative for the other relevant ones
2: by logistic regression method adjusted for repeated observation effect with random effect model
Each slide (except the first & last one for each person) went in to two records; once as the initial result and once as the subsequent one (Section 7.5.2)

7.6.3.2. The effect modification of age

The OR (adjusted for repeated intra-person clustering) between *P. falciparum* in two consecutive surveys in children aged less than one year was 9.3; the corresponding ORs in children aged 1-9 years and more than 9 years were 3.1 and 1.5, respectively. Similar age stratum-specific ORs between *P. falciparum* and *P. malariae* were 11.6, 2 and 1.8, respectively; corresponding ORs for *P. falciparum* and *P. ovale* were 6.9, 2 and 2.7 (Table 7-10).

Table 7-10: The age stratum specific ORs (95% confidence interval) between *P. falciparum* and other species, adjusted for repeated observation

Age group (year)	*P. falciparum*	*P. malariae*	*P. ovale*
<1	9.3(7.6-11.5)	11.6(6.8-20)	6.9(2.7-17.7)
1-9	3.1(2.7-3.6)	2(1.7-2.3)	2(1.4-2.7)
=10	1.5(1.4-1.6)	1.8(1.7-2)	2.7(2.2-3.4)
All	1.9 (1.9-2)	2.7 (2.5-2.9)	3.6 (3-4.4)

These ORs show that the strongest associations were in children aged less than one year old; these associations diminished by increasing age. Nevertheless, the relative age patterns of the species-specific ORs were comparable. These results are compatible with the cross-sectional finding (Section 7.6.2.5)

Stronger positive associations in infancy are compatible with the suggestion that less immunity against *Plasmodium* spp, as a result of fewer exposures to infections in the early years of life, causes higher heterogeneity in infection risks according to mosquito exposure risks and maternal acquired immunity (a detailed explanation can be found in Section 7.6.2.6).

7.6.3.3. Temporal variation

In order to assess seasonal variations of the ORs, the records were categorised in three groups based on the date of their latter surveys. Then the ORs (adjusted for the effect of repeated observations) between *P. falciparum* in the former survey and *Plasmodium* spp in the following one were computed. The ORs between *P. falciparum* in two consecutive surveys in the wet season was 4.3, and the corresponding ORs in the dry-cool and dry-hot seasons were 4.3 and 9.8, respectively. Similar season stratum-specific ORs between *P. falciparum* in the former survey and *P. malariae* in the latter survey were 3.6, 4.1 and 5.5, respectively; corresponding ORs between *P. falciparum* and *P. ovale* were 4, 2.6 and 4.7 (Table 7-11).

Table 7-11: The season specific ORs (95% confidence interval) between *P. falciparum* with itself and with other species in consecutive surveys, adjusted for repeated observations

Season*	*P. falciparum*	*P. malariae*	*P. ovale*
Wet	4.3(3.9-4.6)	3.6(3.2-4.1)	4(3.5-6.4)
Dry-cool	4.3(3.9-4.6)	4.1(3.7-4.5)	2.6(2-3.5)
Dry-hot	9.8(9-10.6)	5.5(4.8-6.2)	4.7(2.8-5.7)
All	1.9(1.8-2)	2.8(2.6-3)	3.9(3.3-4.6)

*Based on the date of the latter survey

These ORs show that *P. falciparum-P. falciparum* associations had a wide temporal variation and its highest magnitude was seen in the dry-hot season. Figure 7-4 illustrated a considerable variation in annual *P. falciparum* frequency, which peaked at the end of wet season and had its lowest prevalence with more or less a stable state in the dry-hot season. In other words, the highest OR was observed in the low transmission period. Based on the discussion in Section 6.3, one could infer that greater ORs in

the dry-hot season were the result of higher heterogeneity in infection risks among people; i.e., a subgroup of people had repeated positive slides in the dry-hot season, when the transmission in the whole population was low. Therefore, either this subgroup had a low clearance rate, i.e., long infection, or they had a high acquisition rate in the dry-hot season when transmission was low in the whole population.

The seasonal variation between *P. falciparum* and *P. ovale* was compatible with the corresponding cross-sectional figures (Section 7.6.2.3). This variation has its highest magnitude in the dry-cool season just after the high transmission season and both species peaks.

Although, the differences between stratum specific ORs of *P. falciparum* and *P. malariae* were low, the likelihood ratio test was significant ($p<0.01$). The temporal variation of OR between *P. falciparum* in the former survey and *P. malariae* in latter one was not consistent with cross-sectional findings (Section 7.6.2.3). Cross-sectional analysis showed that this OR reached its highest magnitude in the dry-cool before *P. malariae* and after *P. falciparum* peaks. In contrast, longitudinal analysis showed the OR had its highest magnitude with delay in the dry-hot season.

These interesting results support the hypothesis that a considerable proportion of *P. malariae* infected subjects in the dry-hot season had probably masked *P. malariae* and patent *P. falciparum* infection in the previous survey. This is also compatible with the proposed explanation about the suppressive effect of *P. falciparum* as the source of delay in a *P. malariae* surge (Sections 4.4.1 and 7.7.6).

The ORs derived in cross-sectional and longitudinal analyses were different. The most reasonable explanation might be the difference in the definitions of dually infected subjects in these two types of analysis. In the cross-sectional analysis, the risks of concurrent infections were assessed, but in the longitudinal analysis, the relationships between infections in two consecutive rounds were checked. The former analysis showed a stronger association between *P. falciparum* and *P. malariae* in the dry-cool season. However, longitudinal analysis showed a greater *P. malariae* positive rate in those who, in the previous round of the dry-hot season, were negative for *P. malariae* and positive for *P. falciparum*. Therefore, the suppressive effect was more evident in the cross-sectional OR, while temporal variations of species and cross-immunity were more important in the longitudinal OR (Section 7.7.3)

7.6.3.4. ORs among the same species

The crude and adjusted (for repeated observations) ORs for *P. falciparum* as a risk factor for the same species in a following survey were 6 and 1.9. The corresponding figures were 6.7 and 2.7 for *P. malariae*, and 8.4 and 5.3, for *P. ovale*. Furthermore, the intra-person correlations (rho) of *P. falciparum*, *P. malariae* and *P. ovale* infections were 0.73, 0.33 and 0.17, respectively (Table 7-9). The above results indicate that the OR for *P. falciparum-P. falciparum* was higher than those of *P. falciparum-P. malariae* and *P. falciparum-P. ovale*.

P. falciparum was much more common than the other two species in the Garki area (Figure 7-2), and compared to the other species, its estimated

acquisition rates were greater in all age groups *[138]*. Further, people acquired immunity against *P. falciparum* faster in the early years of their lives. While 100% of babies showed *P. falciparum* antibody around their first birthdays, *P. malariae* antibody was detected in all children after five year age *[211]*. Although these antibodies are not protective on their own, in endemic areas they increases with age, and are usually taken as a measure of immunity *[92]*.

Most available evidence points to a major role of CD4+ T cells in controlling blood stages of *Plasmodium* spp. *[92,212-214]* The antigen-antibody complex has a major role in activating CD4+ T cells both directly as well as via the complement pathway *[91,213-215]*. Furthermore, in vitro studies show that antibody binds to antigens on extracellular stages of malarial parasites or infected erythrocytes, enabling recognition of the parasite by neutrophils and macrophages via Fc receptors *[216]*.

P. falciparum has the highest blood density among *Plasmodium* spp (Table 7-5); therefore, the immune system is more exposed with *P. falciparum* antigens. For that reason, all types of immune cells (particularly CD4+ T cells, neutrophils and macrophages) might eliminate the blood stages of *P. falciparum* faster than those of other species, especially in those who acquired effective immunity due to previous infection.

According to the above explanation, the lower observed ORs in *P. falciparum-P. falciparum* (1.9) compared to the other species (2.7 in *P. malariae-P. malariae* and 5.3 in *P. ovale-P. ovale*) might be explained as follows. Humoral and cellular immunity helped to efficiently eliminate parasites in subjects with a higher exposure to *P. falciparum* (Figure 7-3),

especially when compared to the other species (also Section 7.7.4, and above paragraphs). In addition, higher *P. falciparum* density (as a result of recent infection) exposed immune cells to a high load of antibody-antigen complex, which in turn stimulated immune cells to eliminate *P. falciparum* faster than other species. This explanation is also supported by Molineaux and Gramiccia (1980) *[138],* who found that the daily clearance rate of *P. falciparum* was significantly lower than those of the other species in early years of life; however, among adults, the differences were not significant *[138].*

Even though *P. malariae* has a lower blood density than *P. ovale,* (Table 7-5) it produced a greater OR (see above paragraphs), which was not compatible with the *P. falciparum* pattern described above. This finding could be explained by a lower prevalence of *P. ovale* in the Garki area. Despite a higher load of *P. ovale* antigens (due to recent infection), a history of less exposure to *P. ovale* infections meant that immune cells might not have reacted quickly enough to eliminate *P. ovale* parasites as rapidly as they eliminated those of *P. malariae.*

A higher rho of *P. falciparum* meant stronger intra-person clustering; i.e., having known a subject's blood slide results, the prediction of infections in their following blood slides would be more accurate in *P. falciparum* compared to the other species. In other words, a person infected with *P. falciparum* was more likely to have a positive slide in the following surveys, (although not necessarily the next one because of a lower *P. falciparum* OR) and negative people had less chance of having a positive slide during their follow up.

A plausible explanation for the above finding is as follows. Acquired immunity against *P. falciparum* increased its clearance rate, as explained in the above paragraphs. However, much stronger immunity might have also decreased the acquisition rate. Therefore, people with very high immunity against *P. falciparum* showed few episodes of infections, while semi-immune people got infections but eliminated *P. falciparum* faster than they did other species. This explanation is also supported by lower age of infections with *P. falciparum* (section 7.6.2.5) because older people acquired immunity against *P. falciparum* infection. However, since the mean ages of infection with *P. malariae* and *P. ovale* were greater than the mean age in the whole population, people were infected with both species during their whole lives.

7.6.3.5. ORs among different species

The adjusted ORs between *P. falciparum* in the former survey and *P. malariae* and *P. ovale* in the latter one were 2.68 and 3.6, respectively (Section 7.6.3.1). Alternatively when exposure and risk were reversed, the adjusted ORs between *P. malariae* and *P. ovale* in the former survey and *P. falciparum* in the following round were 1.7 and 1.91, respectively (Table 7-9).

The OR for *P. falciparum* as risk factor of *P. malariae* was less than the corresponding OR between *P. falciparum* and *P. ovale,* but greater than that for *P. falciparum* and *P. falciparum,* whose OR was 1.9. However, the reverse ORs for *P. malariae* and *P. ovale* as risk factors of *P. falciparum* are more or less comparable (1.7 and 1.91, respectively).

Although, all the ORs are greater than one, possibly due to the heterogeneity effect in infection risks (Section 7.7.2), ORs closer to one could be inferred as either stronger cross-immunity or stronger suppressive effects between species, respectively explained as follows:

1. *P. falciparum* had stronger cross-immunity, i.e., a stronger tendency to prevent new infections, against *P. malariae* (OR=2.7) than it did against *P. ovale* (OR=3.6). However, the reverse cross-immunities were more or less the same (OR=1.9 and 2.1 for *P. malariae* and *P. ovale*, respectively); i.e., immunity against *P. malariae* and *P. ovale* roughly had a similar impact on *P. falciparum*.
2. There was a stronger *P. falciparum* suppressive effect on *P. malariae* than *P. ovale*. This implies that people with mixed infections of *P. falciparum* and *P. malariae* did not show their *P. malariae* infection for longer time (even after clearance of *P. falciparum*) compared to people with mixed *P. falciparum* and *P. ovale* infections.

Comparing the effects of *P. falciparum* on the other species with their reverse effects on *P. falciparum* shows that other species had greater impact on subsequent *P. falciparum* infection, and shifted the ORs closer to one. The adjustment for potential confounding effects of age and season did not change this ordering. To explain this finding, two hypotheses can be proposed:

1. Other species generated stronger cross-immunity against *P. falciparum* than *P. falciparum* generated against the others. However, this explanation is not compatible with previous findings;

i.e., *P. falciparum* was usually the dominant species, had higher density and stimulated the immune system more effectively (Section 7.6.3.4).

2. *P. falciparum* infection suppressed other species below patency but it did not eliminate them. Therefore, some of the *P. falciparum* positive group (with undetectable mixed infections) in the former survey showed other, previously concealed, infections in the following round as the force of *P. falciparum* declined. This explanation is also supported by other finding (see above paragraphs and Section 7.7.6).

7.6.3.6. The conversion rates

Based on the transition numbers from positive to negative and reverse figures, the daily acquisition and clearance rates were computed and classified by age group (Section 7.5.2, and Appendix 3). Checking the effect of *P. falciparum*, the conversion rates of other *Plasmodium* spp were computed based on the presence or absence of *P. falciparum* in the former survey. These methods are similar to those in used by Molineaux and Gramiccia *[138]*, except he did not assess the *P. falciparum* effect on the conversion rates of the other species.

The applied formulae are valid only in stable situations, i.e., situations with constant *Plasmodium* spp prevalence. To assess the validity of the computations, the estimated frequencies were compared with their corresponding observed frequencies. The computed frequencies based on these conversion rates were very close to the observed frequencies (less than

1% difference), except for the *P. falciparum* frequencies in less than one year-old babies (observed 52% and predicted 69%). It seems that the main reason for these discrepancies was due to seasonality of infections. Therefore, the estimated conversion rates should be explained with care.

In both *P. malariae* and *P. ovale*, the highest acquisition and lowest clearance rates were seen in the middle age group, children aged 1-9 years old. This means that 1-9 year old children were more susceptible to infections, and they were infected for longer periods of time.

In addition, the conversion rates were greater in *P. falciparum* positives in all age groups and both species. Clearance rates were lower in *P. falciparum* positives and the age patterns were the same in two species. The differences between clearance rates of *P. falciparum* positives and negatives were more prominent in the older age group in both species.

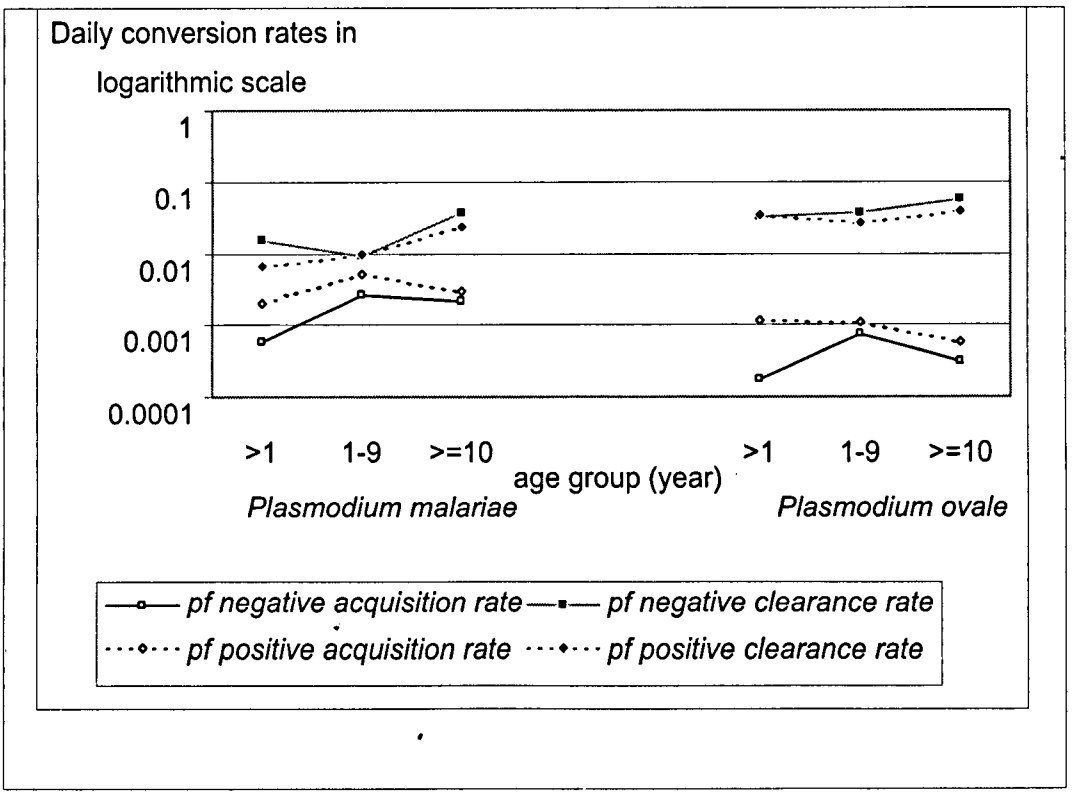

Figure 7-6: Estimated daily clearance and acquisition rates of *P. malariae* and *P. ovale* classified by the presence of *P. falciparum* in the former survey.

The observed differences in acquisition rates might be explained by either:

1. The association in the exposure risks, which means that given *P. falciparum* infection, the person had a higher risk of effective contact with a vector, thereby contracting every species.

2. The suppressive force of *P. falciparum* on the other species, which means that some of the *P. falciparum* positives in the former survey had other infections that were non-detectable by microscopy diagnostic tools. However, by the time the force of *P. falciparum* diminished, the hidden infections were detected.

Lower clearance rates in the *P. falciparum* positives can also be explained by the following suggestions:

1. Higher super-infection rates in *P. falciparum* positives, which might have been due to more exposure to vectors
2. Higher susceptibility to all *Plasmodium* spp in *P. falciparum* positives. Since the acquisition rates in all age groups were higher in *P. falciparum* positives, susceptibility due to other causes such as genetic factors or other illnesses should also be taken into account.

In summary, *P. falciparum* infection decreased the parasite clearance rate; i.e., *P. falciparum* positive subjects had longer infection periods. In addition, *P. falciparum* positive subjects had greater acquisition rates.

7.6.3.7. Summary of the longitudinal results

In the longitudinal approach, the ORs between *Plasmodium* spp in two consecutive surveys were assessed, and the results were shown in Table 7-9. Section 7.6.3.1 presented the ORs between *P. falciparum* in the former survey on the other species in the latter survey; the ORs between *P. falciparum* and *P. malariae,* and *P. falciparum* and *P. ovale* were 2.7 and 3.6, respectively. These results suggested that *P. falciparum* had stronger cross-immunity with *P. malariae* than with *P. ovale,* or *P. falciparum* suppressed *P. malariae* more than it suppressed *P. ovale.*

The age stratum-specific ORs showed that the strongest associations were in the children less than one year old; by increasing age, these associations

diminished (Section 7.6.3.2). These finding were compatible with suggested interactions between exposure risks and acquired immunity as the main source of heterogeneity in infection risk, as discussed in Section 7.6.2.6.

These ORs had seasonal variations, and reached their highest magnitudes in the dry-hot season. Section 7.6.3.3 explored the possible explanations for these temporal variations, which supported the hypothesis of *P. falciparum* having suppressive effects on the other species, particularly *P. malariae.*

The OR (adjusted for repeated observations) between *P. falciparum-P. falciparum* was 1.9; the corresponding ORs between *P. malariae-P. malariae* and *P. ovalė-P. ovale* were 2.7 and 5.3, respectively. Section 7.6.3.4 discussed the differences between these ORs and concluded that the lower *P. falciparum-P. falciparum* OR might be due to its greater clearance rate as well as the result of stronger acquired immunity.

Section 7.6.3.6 presented the daily acquisition and clearance rates of *P. malariae* and *P. ovale,* classified by age group in *P. falciparum* positive and negative groups. In all groups, the acquisition rates were greater, and the clearance rates were lower in *P. falciparum* infected subjects. Section 7.6.3.6 explained these findings based on either suppression of *P. falciparum* on the other species or higher exposure risk of *P. falciparum* positives.

The following section links the findings from the previous sections and discusses their possible immunological and non-immunological mechanisms.

7.7. Discussion

Within the results of Molineaux and Gramiccia book *[138],* there is a short section about the interactions between species. Using a simple statistical approach, the authors showed positive associations between species, and suggested that the associations were mostly due to a subgroup of the population who had a higher susceptibility to all species (Section 4.3.2).

The Garki project was one of the largest field studies of malaria. In addition, it had a large sample size, and measured a wide range of variables with a long follow-up period. Therefore, even after nearly 30 years, it can be used as a comprehensive epidemiological data source.

This analysis was based on the Garki data to specifically address the interactions between *Plasmodium* spp. In the Molineaux and Gramiccia book, the interactions between species were analysed with a cross-sectional approach and did not assess the effects of age, *P. falciparum* density and seasonal variations. This study explored the associations between *Plasmodium* spp with both cross-sectional and longitudinal approaches. The effects of age, season and temporal variations were also explored.

The results are important not only to explain the differences between cross-sectional and longitudinal approaches but also to explore the possible explanations about the immunological and non-immunological interactions between *Plasmodium* spp.

7.7.1. The accuracy of microscopy results

In the assessment of the frequency of dually infected slides by microscope, observational bias in reading films is an important issue (Section 4.4.5).

In the Garki project, most of the blood slides were assessed for 200 fields, and a systematic random sample of one-fifth was examined for 400 fields. This study analysed only the data of 200 fields; the positive films in 400 fields were allocated to positive or negative by a random experiment based on the number of fields found positive with the same method which had been used in the Molineaux and Gramiccia book *[138]*.

Although it is impossible to rule out the effect of observational bias completely, it seems that the data in the Garki project had been collected in an accurate way. Throughout the duration of the Garki project, a systematic random sample of blood films examined by microscopists had been re-examined by a supervisor and the microscopists had not known which of these films would have been re-examined. In total, the kappa agreement coefficients, between microscopists and supervisors were more than 0.77 for all species *[138]*. This implies acceptable agreement between the findings of microscopists and supervisors.

7.7.2. Positive associations between species

Reviewing the literature shows a wide range of associations from negative to highly positive correlations between *Plasmodium* spp (chapter 5). Chapter 6 explained the impact of heterogeneity in infection risks on the overall OR, and illustrated that the observed positive ORs could be explained just by the confounding effects of the heterogeneity in infection risks.

To explain the source of heterogeneity, three suggestions could be proposed which are not mutually exclusive:

1. Similarity in the transmission routes within *Plasmodium* spp and heterogeneity of exposure risk to vectors among the population
2. Acquired immunity, which means that a group of people with lower immunity will have higher risk for all species. In practice, age is a proxy variable which may show the strength of acquired immunity in highly endemic areas.
3. Heterogeneity in susceptibility to *Plasmodium* spp due to genetic or other factors, such as malnutrition, which means that a group of people is prone to all infections.

The Molineaux and Gramiccia book mostly discussed the importance of the second explanation. Although there is some evidence to support all of the above hypotheses based on the results of the Molineaux and Gramiccia book and this analysis, it seems that the first and second hypotheses together could explain most findings (Sections 7.6.1.3, 7.6.2.3, 7.6.2.5, 7.6.2.8, 7.6.3.2, 7.6.3.3, and 7.6.3.6).

The Molineaux and Gramiccia book showed an excess of persons who tended to be positive or negative in consecutive surveys compared to a binomial distribution with the same average, especially in older age groups. Adjusting for the seasonal effect explained a considerable portion of this skewness *[138]*. They concluded that heterogeneity in the exposure risk to mosquitoes, i.e., the temporal variations of mosquitoes, was an important source of the positive skewness, and was more prominent in the older age group.

This analysis showed that the highest ORs between species by far were in infants in both the cross-sectional and longitudinal analyses (Sections 7.6.2.5 and 7.6.3.2). However, the similarity in exposure risk hypothesis proposed in the Molineaux and Gramiccia book can not easily explain the higher ORs in infants.

To address the very high ORs in infants, Section 7.6.2.6 explained this finding based on the interaction between exposure risks and acquired immunity; i.e., in the older age group a stronger acquired immunity in a highly exposed group might decrease the forces of these two factors as sources of heterogeneity of infection risks. While during the early years of life, even among highly exposed infants, immunity was low.

7.7.3. Cross-sectional versus longitudinal results

Most findings about species interactions come from cross-sectional studies. However, it should be mentioned that the two types of studies measure different things. A cross-sectional study measures the risk of concurrent mixed infections, whereas, a longitudinal study measures the risk of one infection based on the history of another species in the past.

The force of suppressive effects between species is an important issue in cross-sectional analysis; however, its force in longitudinal analysis depends on the gap between consecutive assessments. In addition, the results from longitudinal studies depend on the force of cross immunity between species. Assuming strong suppressive effects between species, one would expect to see stronger negative associations; i.e., smaller ORs in cross-sectional

studies. In other words, the dominant species pushes the other ones' densities to levels undetectable by microscope.

On the other hand, short term acquired immunity and suppression effects may decrease the ORs; i.e., more negative association in longitudinal analysis. Given effective cross-immunity between species, a person in a high risk group for one species would be moved to a low risk group after contracting other species. This means that the risk of infections would have been lower, if one had had an infection in a previous round. Also, the suppressive effect of *P. falciparum* may play an important role in a longitudinal analysis if the gap between two consecutive blood tests is short.

However, it should be added that other transient factors related to infection risk, such as, within personal variation of the exposure risk or concurrent diseases, might also reduce ORs in longitudinal analysis. Section 6.3.3 shows that the OR would be confounded by heterogeneity in infection risks if heterogeneity exists in both species; i.e., even maximum heterogeneity in the risk of one species (maximum k_1) with no heterogeneity in the risk of the other species (k_2=1), does not inflate the overall OR. Therefore, a lower association between species would be expected if a considerable proportion of highly exposed people in the first survey moves to the other group and vice versa. In other words, heterogeneity in exposure status over time is an important factor in only the longitudinal approach. Based on this justification, the effect of cross-immunity between species on longitudinal ORs could be explained from a different point of view: cross-immunity might change the temporal infection risks of individuals since the infection risk of people in the second round would be less if they had contracted another species previously.

In this analysis, only records in non-intervention groups were included, in which people did not have access to any anti-malarial drugs and were mostly asymptomatic *[211]*. Nevertheless, the residual drug effect may partially protect cases against the secondary attacks in the short term, and change the pattern of ORs in longitudinal analysis, as well.

To summarise, it could be suggested that the suppression effect might decrease the positive associations between species in cross-sectional analysis; however, the cross-immunity, suppression and temporal variation of infection risks may decrease the positive associations in the longitudinal analysis.

In this study, the ORs based on longitudinal analysis were less than the corresponding ones on cross-sectional analysis (Sections 7.6.2.2 and 7.6.3.1). This implies that the suppression of one species with concurrent infections had less effect than within personal variations of susceptibility to infection risks due to acquired immunity, concurrent diseases or variation in exposure risks, or the short term suppression effect.

The ratio between ORs of *P. falciparum* with *P. malariae* in longitudinal and cross-sectional analyses was 0.73 (2.68/3.64); the corresponding ratio between *P. falciparum* and *P. ovale* was 0.71 (3.6/5.1). Hence it seems that the pattern of differences in cross-sectional and longitudinal results between *P. falciparum* and other species are comparable.

In conclusion, it should be mentioned that the cross-sectional and longitudinal ORs measure different things and to explain these difference, other factors, such as the gap between two observations, the effect of anti-

malarial drugs, and the temporal and spatial variations of species, should be taken into account.

7.7.4. Acquired immunity

Table 7-5 shows that in naïve people, the duration of infections are very long; i.e., very low daily clearance rates. There is no evidence to support that within susceptible people, infections resolve as a result of intrinsic features of the parasite. For this reason, the clearance rate of infections without any medication might be inferred as the force of immunity against species.

In malaria endemic areas, the levels of antibodies increase with age. *[92]* Although the correlation between antibody levels and protection is low, their levels are usually taken as a measure of exposure, and antibodies play an important role in the clearance of parasites *[92]*.

Serological data of the Garki project also showed that antibodies against *Plasmodium* spp rose with age and had negative correlations with infection risks *[211]*.

Efficient production of anti-malaria antibodies however requires an intact and functioning T cell system *[91,212-215]*. Furthermore, possessing antibodies does not protect against infections unless T cells, particularly CD4+, and other immune system components work functionally *[92,212-214]*. On the other hand, specific T cell immunity and non-specific immunity (via the complement pathway, the neutrophils, macrophages and natural killer cells) work much more efficiently in the presence of antibodies

[216-218]. Thus, it could be inferred that humoral and cellular immunities have synergistic effects on the elimination of parasites.

Section 7.6.3.6 shows a lower clearance rate among of children, suggesting that acquired immunity due to a history of infections plays an important role in the clearance of infections.

However, Molineaux and Gramiccia (1980) showed that older people were also more susceptible to infections in the beginning of the transmission season *[138]*. Although older people were prone to infections in the beginning of the wet season, their *P. falciparum* densities were much lower than the level of parasitaemia in young subjects (Section 7.6.2.6). Therefore, it seems that long-term acquired immunity reduces the severity of infections and increases the clearance rate.

Based on the above finding, it seems that a frequent booster effect due to natural infections, or the presence of antigens, is needed to keep up effective immunity against *Plasmodium* spp (see following paragraphs); and even a short period, such as a few months, is enough to reduce the protective force of immunity (Section 7.7.4).

According to available immunological findings, it seems that acquired cellular and humoral immunities keep their functions up for months or even years after frequent exposures. In the secondary immune response, Immunoglobulin G (IgG), which has a long half life, increases. Cellular immunity also has long term effects, as it generates memory cells, which remain in the body for the rest of life *[219,220]*. Therefore, high susceptibility within older people in the beginning of the transmission season is not easily explained by declining immunity in the short term.

Genetic diversity among *Plasmodium* spp and annual mutations in their antigens could be an alternative explanation for higher susceptibility of older people in the beginning of transmission seasons. While, this explanation should not be rejected, *P. falciparum* peaked only for few weeks in the beginning of the transmission season within adults. Based on current knowledge about immunity against malaria, it seems that the immune system needs a longer time to obtain enough strength to eliminate parasites with new antigens *[221-224]*.

The above explanation is also supported with the long duration of infections among children who had low acquired immunity (Sections 7.6.2.5, 7.6.3.2, and 7.6.3.6). In other words, it seems that a few weeks is not enough time for an immune system to cope with new antigens even in those who had memory against similar antigens.

In addition, mutations are more likely to occur during transmission seasons, when *Plasmodium* spp proliferate faster. In this case, some peaks of infections among adults at the end of transmission seasons would be expected; but this hypothesis was not supported with the Garki data.

Alternatively, it is known that within people who have activated immune cells and antibodies, the presence of the antigen-antibody complex improves the efficiency of immune systems considerably *[91,92,212-214,216]*. It could be suggested that in the beginning of transmission season, adults showed lower protection against *Plasmodium* spp due to the absence of antigens, and they contracted infections, more or less at the same rate as children. However, soon afterward, the presence of antigens strongly stimulated the adult immune systems, resulted in higher clearance rates in adults than

children. In other words, the complex of antigen-antibody might magnify the specific T cell functions and non-specific immunity.

Based on this suggestion, even an immune system which is activated due to a history of exposures has a delay in responding against *Plasmodium* spp functionally. This delay cannot be explained by humoral immunity because immunoglobulin has a long half-life in serum and attaches rapidly to the antigens *[219,220]*. However, cellular immunity might need more time, since it needs the complex of antigen-antibody, processed antigens by Antigen-Presenting Cells (APC) conjugated with Major Histocompatibility Complex (MHC) class II proteins, and a wide range of cytokines *[225-231]*.

In summary, it could be inferred that, although long term acquired immunity plays an important role against infections, it needs to be stimulated by an antigen-antibody complex and processed antigen epitopes.

Some longitudinal studies in Irian-Jaya compared the prevalence of parasitaemia in native people with trans-migrants from an area of low malaria transmission. In the first year of exposure, the risks of infection in trans-migrants were higher than in native people in all age groups. After 14 months however, children showed a higher density of parasites, and malaria-naïve adults acquired partial resistance to infection faster than naïve children. Based on this finding, writers concluded that protective immunity against *P. falciparum* developed after only a relatively brief period of exposure, and the efficacy of protection was profoundly affected by a function of age unrelated to cumulative exposure. In other words, the degree of protection was governed by recent exposure and age, independent of history of chronic heavy exposure *[221-224,232]*.

Although intrinsic features of the immune system that change with age may determine key characteristics of the immune response to the infection, non-specific immunity, i.e., activated T cells even with non specific antigens, might also have a role *[221]*. The interaction between age and acquired immunity might be explained by the higher non-specific immunity in adults, perhaps acquired via other infections.

Compared to other species, *P. falciparum* had the highest clearance rate based on the ORs between two consecutive positive slides (Sections 7.6.3.4 and 7.6.3.5). This implies that *P. falciparum* stimulated the immune system more efficiently than other species. Section 7.6.3.4 explains this finding based on greater exposure to *P. falciparum* in the past and a higher density of *P. falciparum* as a result of recent infections.

7.7.5. Cross-immunity between species

There is little published information on the cross-immunity between *Plasmodium* spp and most of it is from experimental studies in animals which showed that cross-immunity between species did not give considerable protection *[90,142]*. Bruce et al. (2000) showed that *Plasmodium* spp are independent in humans *[134]*. Section 7.6.3.7 explains the findings regarding the ORs between *Plasmodium* spp. Based on these results, it is very difficult to differentiate the impact of cross-immunity between *Plasmodium* spp and the short-term suppressive effect. The protective effect of the acquired long-term immunity against each species was low, which goes against the cross-immunity hypothesis.

According to the presented results, it seems that cross-immunity did not have any considerable impact on the burden of infections even in older people. However, the role of non-specific immunity should be taken into account (Section 7.7.4). This means that even without specific cross immunity between species, infection with one species activates non-specific immunity, which might boost the strength of immunity against other species as well.

7.7.6. Suppressive effect of *P. falciparum*

It is not possible to differentiate the cross-immunity and suppressive effects between *Plasmodium* spp. The suppressive effect may be mediated by competition for host cells or nutrients, or by heterologous immunity, which is the same mechanism for cross-immunity. To be more specific about the role of immunity in these two effects, it can be said that cross-immunity protects subjects against the acquisition of a second infection or increases its clearance rate, while the suppressive effect just prevents the parasitaemia of the second species for a short term while the dominant species has its peak.

In agreement with others findings, the results of this analysis showed that *P. falciparum* suppressed the expression of other species. Looareesuwan et al. (1987) *[135]* found considerable *P. vivax* relapses, more than expected by chance, after treatment of *P. falciparum* cases and explained these findings by the suppressive effect between *Plasmodium* spp. The suppression hypothesis is supported by data derived from the simultaneous inoculation of two *Plasmodium* spp into laboratory animals; many studies

have shown that one or both species are suppressed *[136,140]*. However, the suppressed species rebounds after the other species has abated, and may show a prolonged infection *[136,140]*.

There is consistency between the cross-sectional and longitudinal results in this analysis about the suppressive effect of *P. falciparum* on the other species, particularly *P. malariae.*

P. malariae peaked in the dry-hot season around 6 months after the *P. falciparum* peak in the transmission season (Section 7.6.1.3). However, the *P. malariae* infant conversion rate had a strong positive association with Vectorial capacity *[138]*. Longitudinal analysis showed that a considerable proportion of *P. malariae* positives in the dry-hot season had a positive *P. falciparum* slide in their previous survey (Section 7.6.3.3). In addition, mixed infected slides were detected mostly after the wet season, which again, can be explained by the suppressive effect of *P. falciparum* as the dominant species. That is, as *P. falciparum* density declined following the transmission season, the chance to become patent increased in other previously masked species (Section 7.6.2.3).

Section 7.6.3.5 shows that the ORs between *P. falciparum* in the former survey and other species in the latter survey are greater than their reverse temporal order. Because of the non significant cross-immunity impact, the protective effects of other species against *P. falciparum* are particularly questionable. The suppressive effect of *P. falciparum* on the other species in the former survey seems the only plausible explanation (Section 4.4.1).

Furthermore, the suppressive effect of *P. falciparum* can explain the higher daily acquisition rates of *P. malariae* and *P. ovale* in *P. falciparum* positive people (Section 7.6.3.6).

Table 7-2, which summarizes the impact of interventions in the Garki project, showed that in contrast to *P. falciparum* and *P. ovale, P. malariae* frequency did not drop by using insecticide in the short term. This finding could be also explained by the suppressive effect of *P. falciparum* as follows. The *P. falciparum* incidence decreased as the result of insecticide impact on the Vectorial capacity. In turn, the suppressive force of *P. falciparum* on the other species decreased and a reverse surge in the *P. malariae* incidence was observed.

The risks of *P. malariae* and *P. ovale* had positive associations with the *P. falciparum* density. If there had been competition, it would be expected that a high density of *P. falciparum* infection would give other species less chance of expression.

Alternatively, the suppression phenomenon could be considered as a type of cross-immunity which does not eliminate other species, but pushes their densities to undetectable levels. Thus, it can be concluded that *P. falciparum* has a considerable suppressive effect on the other species, particularly *P. malariae*, and that this effect is mediated by immunological inferences rather than competition for host cells or nutrients. This effect depends on previous exposure experiences. The effect fades just after the *P. falciparum* peak. This suggests that the complex of antigen-antibody may inhibit the proliferation of other species. This hypothesis can be used to

explain the differences between interactions in all age groups and their temporal variation.

7.7.7. Addressing the study objectives

The main findings in this study showed very strong positive associations between the risks of *Plasmodium* spp infections in both cross-sectional and longitudinal analyses. However, the associations were stronger in the cross-sectional analysis. The intra-person clustering effect was stronger for *P. falciparum* infection and all types of ORs were greater by far in infants. ORs between species were significantly affected by age, *P. falciparum* blood density, season and location.

These findings, compatible with findings in other studies, can be explained by the heterogeneity in infection risks and their temporal and spatial variations, the suppressive effect of *P. falciparum* on the other species and immunological pathways. In addition, it can be implied that immunological mechanisms had an important role on the suppressive effects of *P. falciparum* on the other species.

CHAPTER 8

8. Overall discussion

This chapter reviews the main findings and expresses the final conclusions. It concludes by formulating initiatives for further research.

8.1. The relationship between Part One and Two

This thesis explored two aspects of malaria, linked by the need to understand the differing epidemiology of *Plasmodium* spp. The first part assessed the feasibility of an early warning system based on climate and remote sensing data. The second part was about the interaction between *Plasmodium* spp.

The spatial and temporal variations of *Plasmodium* spp were important in both parts. The first part, assessed whether the temporal and spatial discrepancies between species could decrease the accuracy of the predictions. The second part discussed whether these discrepancies were entirely due to the interaction between species.

Acceding to the results these, it seems that species-specific models would improve the predictions due to the different impacts of climate on the transmission of species and the interaction between them. In addition, the

temporal and spatial spans of studies may change the overall pattern of the interactions between species.

8.2. Feasibility of prediction models

It is well known that climate affects malaria epidemics. Also, there has been a great deal of research on the modelling of malaria based on meteorological factors. However, most of these studies examined the accuracy of their models based on malaria data in epidemic areas mostly in Africa. In addition, Geographical Information System (GIS) and Remote Sensing (RS) data has not been used in this field until recently.

For these reasons, the decision to assess the feasibility of an early warning system in an endemic area in Iran with seasonal malaria using GIS and RS data was made.

The main goal in this part of thesis was to evaluate the accuracies of the models and assess their practical applications in the field. It also focused on the scientific aspects of the models and introduced some new epidemiological methods.

The selected area, Kahnooj district, had an acceptable surveillance system and its health system cooperated fully to provide all required data. In addition, its ground climate data has been collected in a standard format and archived systematically for the past 30 years.

8.2.1. Main findings

These analyses assessed the feasibility of an early warning system and provided an in-depth exploration of the epidemiology of malaria in Kahnooj in terms of the malaria risk factors, the seasonal variations, *P. vivax* relapse rate and therapeutic failure rate.

To address the main objectives of this part of the thesis, accuracy of models were evaluated by comparing the model predictions to the observed numbers of malaria cases in the checking part of the data.

According to the findings, it seems that *P. falciparum* and *P. vivax* have different sensitivities to the meteorological factors, and that these differences explain most of the discrepancies in their temporal variations. Hence, it is suggested that distinct species-specific models based upon climate fit the variations more precisely.

The findings from this part of the thesis also imply that the models based on the ground climate data were more appropriate than the models based on the remote sensing data. However, for final conclusions, more studies are needed to assess the accuracies of models with remote sensing data in finer scale and with more data, particularly rainfall estimation.

The models predicted the number of malaria cases one month ahead, which is enough for the health system to reinforce its control programmes in high risk areas.

The models also predicted the number of cases at three distinct spatial levels: district, subsubdistrict (SSD) and village levels. The accuracies of these models had a negative relationship with their spatial levels, i.e., the

models at the district and village levels had maximum and minimum accuracies, respectively.

The accuracies of the predictions at the district level were improved by using the ground climate data. Although predictions at the district level may help to establish an early warning system to improve the effectiveness of control programmes, from a practical point of view, it seems that predictions at the SSD level would be more appropriate.

In this study, ground climate data were available only at the district level from the synoptic centre in Kahnooj city, which may be one of the main reasons for the low accuracies observed in the SSD and village models.

Measuring the required meteorological factors for these models is very simple and does not need any special expensive equipment or highly skilled personnel. Climatological centres are located in different parts of the district and measure the required meteorological factors continuously. Nonetheless, the reporting system to district and province centres is not as well-established as the reporting system of the synoptic centre. Therefore, it seems that with a small effort, the system could be improved to provide accurate and up-to-data climate data, even at the SSD level, which may improve the accuracy of the SSD model considerably.

As final conclusion it seems that the model based on ground climate data is feasible to predict the number of cases one month ahead in the district level.

8.2.2. Research initiatives for further studies

1. This study did not quantify the feasibility of the predictions in terms of cost-effectiveness. Further cost-effectiveness analysis could provide much more objective measures to judge the feasibility of the models.
2. RS data with finer resolution and with greater number of variables such as rainfall and different vegetation indices may improve the model accuracies. Therefore, further studies in this field may help to clarify the feasibility of the RS data in the prediction of malaria epidemics.
3. It may be worth to assess the feasibility of the models based on extrapolated meteorological data extracted from the data of synoptic centres around the Kahnooj district for every village and SSD. These models may show whether synoptic data are enough to improve the accuracies of the SSD models or whether meteorological measurements from each SSD are also needed.

8.3. The interactions between species

The second part of this thesis assessed the interactions between species from different points of view: systematic review of the published literature and meta-analysis, modelling, and extended analyses of the Garki data. New hypotheses, compatible with the current knowledge about the interaction between species, were also generated and tested.

The interaction between species is highly dependant on the type of study and many environmental and biological factors which introduce a wide heterogeneity in the interactions between species. Exploring this issue can expand our knowledge about the immunology and epidemiology of malaria. This thesis discussed the possible impacts of acquired immunity and cross immunity and suppressive effects between *Plasmodium* spp.

8.3.1. Main findings

The meta-analysis showed an overall negative association between *P. falciparum* and *P. vivax*. However, a very wide range of associations were observed among studies. This heterogeneity was explained partly by the prevalence of infections and temporal span of studies.

The modelling chapter showed that within-population heterogeneity in infection risk (due to heterogeneity either in the exposure risks or in the susceptibility to infections) may distort the observed OR and can, by itself, explain ORs even as great as ten, which have been observed in some studies.

The Garki data were analysed with two approaches, cross-sectional and longitudinal. The ORs in the longitudinal analysis were less than in the cross-sectional analysis. ORs were also found to be dependent on the age group of subjects, season and *P. falciparum* densities.

In conclusion, an explanation compatible with all the findings in other studies seems to necessitate an important role for acquired immunity, possibly via the combination of the humoral and cellular pathways. Nevertheless, the immune system needs constant exposure to antigens to

maintain effectiveness against *Plasmodium* infections. It seems that the cross-immunity between *Plasmodium* spp does not protect against infections, although the role of unspecific immunity cannot be ruled out.

The results also suggest that the suppressive effect of *P. falciparum* on the other species, particularly *P. malariae*, is considerable. It other words, *P. falciparum* may push the densities of the other species to levels undetectable by microscopy when it has its blood stage peak, but the suppressed species reappear in the blood just after the *P. falciparum* surge. This suppressive effect cannot be explained entirely by the competition between species for red blood cells or nutrients. Therefore, it seems that other immunological pathways, such as non-specific immunity, may decrease the blood densities of other species when the concentration of the complex between antibodies and *P. falciparum* antigens is high.

8.3.2. Research initiatives for further studies

1. The meta-analysis chapter focused on the association between *P. falciparum* and *P. vivax*. To clarify the associations between other *Plasmodium* spp, a systematic literature review and meta-analysis is necessary.

2. Although there are clear findings about the suppressive effect of *P. falciparum* on the other species, the current knowledge of possible mechanisms is poor. It seems that more attention to this area, particularly its molecular aspects, may add valuable information to our knowledge about the biology of *Plasmodium* spp in the human body and the reaction of the immune system.

3. Although this study tried to explore sources of the heterogeneity in the interactions between species, further studies are needed to investigate other possible aspects of this issue, such as the role of specific and non-specific immunity.

4. Vaccine against one *Plasmodium* species may indirectly change patterns and distributions of the others. These effects may be partly due to the interaction between species, although other factors, such as health seeking behaviour, are also important. This could be a valuable topic for further studies to evaluate and model the possible impact of such a vaccine in the global burden of malaria.

Appendices

1. Checklists of reviewing papers in meta-analysis

1. 1. Abstracts

Citation identification number:

1. What is the main object of the study?
   - ❑ Evaluation of the interactions between different *Plasmodium* infections in humans
   - ❑ Estimation of the prevalence/incidence of malaria
   - ❑ Evaluation of the effect of control programs (bed-nets, vaccine...)
   - ❑ Assessment of the efficacy of malaria treatment
   - ❑ Evaluation of the new techniques of diagnosis of malaria
   - ❑ Others:

2. Did they choose random sample of cases?
   - ❑ Yes
   - ❑ No
   - ❑ Not mentioned
   - ❑ Not relevant

3. Did they exclude immigrant cases?
   - ❑ Yes
   - ❑ No
   - ❑ Not mentioned
   - ❑ Not relevant

4. Is the location of study mentioned in the abstract?
   - ❑ Yes
   - ❑ Yes, but not clearly
   - ❑ No
   - ❑ Not relevant

5. Did they calculate the incidence/prevalence of malaria?
   - ❑ Yes
   - ❑ No
   - ❑ Not clearly stated in the abstract

6. Did they state to the frequency of mixed infections in the abstract?
   - ☐ Yes
   - ☐ No
   - ☐ No, but they mentioned that they calculate them

7. Does it have a cross sectional component?
   - ☐ Yes
   - ☐ No

Final conclusion

This paper

1. is not relevant to the aim of meta-analysis[1]
2. does not have the required information
3. for the final decision the full text of the paper is needed
4. is suitable for this meta-analysis

---

[1] Papers that were not considered to be relevant were those in which:

1. The objectives were not relevant to the aim of the meta-analysis (options 3,4 or 5 in the first question, or option 2 of the seventh question)
2. The cases were chosen non-randomly (option 2 in the second question)
3. The immigrant cases were not excluded (option 2 in the third question)
4. The frequency of malaria and mixed infections were not stated in the paper (option 2 in the fifth and sixth questions).

1.2. Full text

General information about the paper

ID number of paper: Name of first author: Date of paper:
Date of data collection:
Type of publication: journal report
Age group:

Methodological issues

The main object of paper
assessment of the incidence/prevalence of malaria
others
Location: Africa Asia America Australia Europe
Name of country: Other geographical information:
Name of area:
Type of climate:
Altitude:
TREATMENT STRATEGY:

Epidemiology of malaria in the region

Stable or endemic malaria
Hypoendemic
Mesoendemic
Hyperendemic
Holoendemic

Duration of study (temporal distribution)

Area of study (spatial distribution)

Unstable or epidemic malaria

Results

Number of examined blood slides:

The number of positive slides for:
falciparum
vivax
Mixed

2. The computation of daily conversion rates

For each species, transition frequencies between consecutive surveys are *N*++, *N*+ -, *N*- + and *N*- -; where N+ + is the number of persons positive at both surveys, *N*+- the number positive at the first survey and negative at the second one, etc.

The *h* and *r* are the daily acquisition and clearance rates from detectable parasitaemia, s_h^2 and s_r^2 their variances, and t the average of days between two consecutive surveys.

$$h = \left(\frac{\alpha}{t(\alpha+\beta)}\right)\ln\left(\frac{1}{1-(\alpha+\beta)}\right)$$

$$r = \left(\frac{\beta}{t(\alpha+\beta)}\right)\ln\left(\frac{1}{1-(\alpha+\beta)}\right)$$

where:

$$\alpha = \frac{N-+}{N-}$$

$$\beta = \frac{N+-}{N+}$$

$$s_h^2 = \frac{(S_1(\alpha V+\beta U)^2 + S_2(\alpha U-\alpha V)^2}{\gamma^4}$$

$$s_r^2 = \frac{(S_2(\beta V+\alpha U)^2 + S_1(\beta U-\beta V)^2}{\gamma^4}$$

where

$$\gamma = \alpha+\beta$$

$$U = \frac{(-\ln(1-\gamma)}{t}$$

$$V = \frac{\gamma}{(1-\gamma)t}$$

$$S_1 = \frac{\alpha(1-\alpha)}{N-}$$

$$S_2 = \frac{\beta(1-\beta)}{N+}$$

3. Multi-level model Structure

The response variables were binary (whether or not the slide was positive or negative for species of interest). A variance components model was fitted to assess the average of Smear Positive Rate (SPR) along with between and within village variations.

I made a 3 level model; survey, person and villages were defined as the first, second and third level respectively, and were indicated by subscript of *i, j* and *k*.

p_{ijk} is the probability of infection in blood slide of *j*th subject in *i*th survey in *k*th village. This probability is as a function of the intercepts as follows:

$$\text{logit}\,(p_{ijk}) = ß_{1jk}x_1, \text{ where } ß_{1jk} = ß_1 + v_{1k} + u_{1jk}$$

In the above equation, x_1 is a constant, and takes the value 1 for all slides. (Subscript *"i"* indicates the first level of variation and is explained in the following paragraphs). The coefficient $ß_{1jk}$ indicates that the intercepts are modelled in this relationship as random at the second and third levels, x_1 was used for this purpose since x_0 was used to specify the variation at level one (survey). Logit is the link function in this model.

$ß_1$ is the fix term in second and third level and indicates the logit of SPR

v_{1k} is the random term in the third level which shows between villages variation of SPR, and has a normal distribution with 0 mean and Ω_v variance.

u_{1jk} is the random term in the second level which shows within village variance of SPR, i.e., between person variations, with a normal distribution with 0 mean and Ω_u variance.

The full model can be written as:

$$y_{ijk} = p_{ijk} + e_{0ijk}\, x_0^*$$

where y_{ijk} are the observed (0, 1) responses, x_0* is equal to 1, and e_{0ijk} is level one random term which shows the level one residuals. The standard assumption is that the response y_{ijk} is distributed as binominal (1, p_{ijk}). It could be written this distribution assumption in a general form as:

$$y_{ijk} \sim Binomial\,(n_{ijk},\ p_{ijk})$$

Where in this case the n_{ijk} are all equal to 1, so that the variance of e_{0ijk} is $p_{ijk}(1-p_{ijk})$

The whole structure of the model is as follows:

$$\left.\begin{aligned} y_{ijk} &\sim \text{Binomial}(n_{ijk},\ \pi_{ijk}) \\ y_{ijk} &= \pi_{ijk} + e_{0ijk} x^{*}_{0} \end{aligned}\right\}$$

$$\text{logit}(\pi_{ijk}) = \beta_{1jk} x_1$$

$$\beta_{1jk} = \beta_1 + v_{1k} + u_{1jk}$$

$$\begin{bmatrix} v_{1k} \end{bmatrix} \sim N(0,\ \Omega_v) : \Omega_v = \begin{bmatrix} \sigma^2_{v1} \end{bmatrix}$$

$$\begin{bmatrix} u_{1jk} \end{bmatrix} \sim N(0,\ \Omega_u) : \Omega_u = \begin{bmatrix} \sigma^2_{u1} \end{bmatrix}$$

$$x^{*}_{0} = x_0 [\pi_{ijk}(1 - \pi_{ijk})/n_{ijk}]^{0.5}$$

$$\begin{bmatrix} e_{0ijk} \end{bmatrix} \sim (0,\ \Omega_e) : \Omega_e = \begin{bmatrix} 1 \end{bmatrix}$$

Bibliography

[1] Gilles HM. The malaria parasites. In: Gilles HM, warrel DA, editors. Essential malariology. London: Edward Arnold, 1993. pp. 12-34.

[2] Malaria (*Plasmodium*) and Other Haemospororina (*Sporoza*). In: Medical and Veterinary Entomology, editor. Kettle DS. Oxon, UK: CAB International, 1995. pp. 558-90.

[3] Kovats RS, Campbell-Lendrum DH, McMichael AJ, Woodward A, Cox JS. Early effects of climate change: do they include changes in vector- borne disease? Philos Trans R Soc Lond B Biol Sci 2001;356 (1411):1057-68.

[4] Ndiaye O, Hesran JY, Etard JF, Diallo A, Simondon F, Ward MN, Robert V. Climate variability and number of deaths attributable to malaria in the Niakhar area, Senegal, from 1984 to 1996. Sante 2001;11 (1):25-33.

[5] Craig MH, Snow RW, le Sueur D. A climate-based distribution model of malaria transmission in sub-Saharan Africa. Parasitol Today 1999;15 (3):105-11.

[6] Reiter P. Climate change and mosquito-borne disease. Environ Health Perspect 2001;109 Suppl 1:141-61.

[7] Bouma MJ, Sondorp HE, van der Kaay HJ. Health and climate change. Lancet 1994;343 (8892):302.

[8] Martens WJ, Niessen LW, Rotmans J, Jetten TH, McMichael AJ. Potential impact of global climate change on malaria risk. Environ Health Perspect 1995;103 (5):458-64.

[9] Zheng KS. To explore the principal vector of malaria by using vectorial capacity. Zhonghua Liu Xing Bing Xue Za Zhi 1989;10 (3):161-3.

[10] Garrett JC. The human blood index of malaria vectors in relation to epidemiological assessment. Bull World Health Organ 1964;30:241-61.

[11] Dye C, Reiter P. Climate change and malaria: temperatures without fevers? Science 2000;289 (5485):1697-8.

[12] Patz JA, Lindsay SW. New challenges, new tools: the impact of climate change on infectious diseases. Curr Opin Microbiol 1999;2 (4):445-51.

[13] Bayoh MN, Lindsay SW. Effect of temperature on the development of the aquatic stages of *Anopheles gambiae sensu stricto* (Diptera: Culicidae). Bull Entomol Res 2003;93 (5):375-81.

[14] Lindsay SW, Birley MH. Climate change and malaria transmission. Ann Trop Med Parasitol 1996;90 (6):573-88.

[15] Gillett JD. Direct and indirect influences of temperature on the transmission of parasites from insects to man. In: Taylor AE, Muller R, editors. The effect of meteorological factors upon parasites. Symposium of the British Society for Parasitology: Oxford: Blackwell Scientific, 1974. pp. 79-95.

[16] Muirhead-Thomson RC. Mosquito behaviour in relation to malaria transmission and control in tropics. London: Edward Arnold, 1951.

[17] Beaty BJ, Marquardt WC. The biology of vector disease vector. Colorado: University press of Colorado, 1996.

[18] Molineaux L, Gramiccia G. Entomology. In: . editor. The Garki Project, Research on the epidemiology and control of malaria in the Sudan Savanna of west Africa. Geneva: World Health Organization, 1980.

[19] Detinova TS. Age-grouping methods in deptera of medical importance. Geneva: World Health Organization, 1962.

[20] Mahmood F, Reisen WK. Duration of gonotrophic cycles of *Anopheles culicifacies Giles* and *Anopheles stephensi Liston*, with observations on reproductive activity and supervivoship during winter in Punjab Province, Pakistan. Mosquito News 1981;41:41-50.

[21] Lindblade KA. Land use change alters malaria transmission parameters by modifying temperature in a highland area of Uganda. Trop Med Int Health 2000;5 (4):263-74.

[22] Haddow AJ. Measurments of temperature and light in artificial pools with reference to the larval habitat of *Anopheles gambia Giles* and *A. funestus Giles*. Bull Entomol Res 1943;34:89.

[23] Jepson WF, Moutia A, Courtois. The malaria problem in Mauritius, the bionomics of Mauritian *anophelines*. Bull Entomol Res 1943;38:177-208.

[24] Bradley DJ. Human tropical diseases in a changing environment. Ciba Found Symp, 1993. pp. 146-62.

[25] Patz JA, Reisen WK. Immunology, climate change and vector-borne diseases. Trends Immunol 2001;22 (4):171-2.

[26] Lindsay SW, Martens WJ. Malaria in the African highlands: past, present and future. Bull World Health Organ 1998;76 (1):33-45.

[27] Thomson MC, D'Alessandro U, Bennett S, Connor SJ, Langerock P, Jawara M, Todd J, Greenwood BM. Malaria prevalence is inversely related to vector density in The Gambia, West Africa. Trans R Soc Trop Med Hyg 1994;88 (6):638-43.

[28] Kondo H, Seo N, Yasuda T, Hasizume M, Koido Y, Ninomiya N, Yamamoto Y. Post-flood--infectious diseases in Mozambique. Prehospital Disaster Med 2002;17 (3):126-33.

[29] Campanella N. Infectious diseases and natural disasters: the effects of Hurricane Mitch over Villanueva municipal area, Nicaragua. Public Health Rev 1999;27 (4):311-9.

[30] Julvez J, Mouchet J, Michault A, Fouta A, Hamidine M. [The progress of malaria in sahelian eastern Niger. An ecological disaster zone]. Bull Soc Pathol Exot 1997;90 (2):101-4.

[31] Faye O, Gaye O, Fontenille D, Hebrard G, Konate L, Sy N, Herve JP, Toure Y, Diallo S, Molez JF, et al. [Drought and malaria decrease in the Niayes area of Senegal]. Sante 1995;5 (5):299-305.

[32] Saenz R, Bissell RA, Paniagua F. Post-disaster malaria in Costa Rica. Prehospital Disaster Med 1995;10 (3):154-60.

[33] Famine-affected, refugee, and displaced populations: recommendations for public health issues. MMWR Recomm Rep 1992;41 (RR-13):1-76.

[34] Hembree SC. Malaria among the civilian irregular defense group during the Vietnam conflict: an account of a major outbreak. Mil Med 1980;145 (11):751-6.

[35] Patz JA, Epstein PR, Burke TA, Balbus JM. Global climate change and emerging infectious diseases. JAMA 1996;275 (3):217-23.

[36] Malakooti MA, Biomndo K, Shanks GD. Reemergence of epidemic malaria in the highlands of western Kenya. Emerg Infect Dis 1998;4 (4):671-6.

[37] Bouma MJ, Dye C, van der Kaay HJ. *falciparum* malaria and climate change in the northwest frontier province of Pakistan. Am J Trop Med Hyg 1996;55 (2):131-7.

[38] Jambulingam P, Mohapatra SS, Govardhini P, Das LK, Manoharan A, Pani SP, Das PK. Microlevel epidemiological variations in malaria & its implications on control strategy. Indian J Med Res 1991;93:371-8.

[39] Bayoh MN, Thomas CJ, Lindsay SW. Mapping distributions of chromosomal forms of *Anopheles gambiae* in West Africa using climate data. Med Vet Entomol 2001;15 (3):267-74.

[40] Greenwood BM, Pickering H. A malaria control trial using insecticide-treated bed nets and targeted chemoprophylaxis in a rural area of The Gambia, west Africa. 1. A review of the epidemiology and control of malaria in The Gambia, west Africa. Trans R Soc Trop Med Hyg 1993;87 Suppl 2:3-11.

[41] Hendrickx G, Napala A, Slingenbergh JH, De Deken R, Vercruysse J, Rogers DJ. The spatial pattern of trypanosomosis prevalence predicted with the aid of satellite imagery. Parasitology 2000;120 (Pt 2):121-34.

[42] Thomson MC, Connor SJ. Environmental information systems for the control of arthropod vectors of disease. Med Vet Entomol 2000;14 (3):227-44.

[43] Hay SI, Snow RW, Rogers DJ. From predicting mosquito habitat to malaria seasons using remotely sensed data: practice, problems and perspectives. Parasitol Today 1998;14 (8):306-13.

[44] Sipe NG, Dale P. Challenges in using geographic information systems (GIS) to understand and control malaria in Indonesia. Malar J 2003;2 (1):36.

[45] Klinkenberg E, Van Der Hoek W, Amerasinghe FP. A malaria risk analysis in an irrigated area in Sri Lanka. Acta Trop 2004;89 (2):215-25.

[46] Noor AM, Zurovac D, Hay SI, Ochola SA, Snow RW. Defining equity in physical access to clinical services using geographical information systems as part of malaria planning and monitoring in Kenya. Trop Med Int Health 2003;8 (10):917-26.

[47] Hay SI, Omumbo JA, Craig MH, Snow RW. Earth observation, geographic information systems and *Plasmodium falciparum* malaria in sub-Saharan Africa. Adv-Parasitol 2000;47:173-215.

[48] Hay SI, Tucker CJ, Rogers DJ, Packer MJ. Remotely sensed surrogates of meteorological data for the study of the distribution and abundance of arthropod vectors of disease. Ann Trop Med Parasitol 1996;90 (1):1-19.

[49] Beck LR, Rodriguez MH, Dister SW, Rodriguez AD, Washino RK, Roberts DR, Spanner MA. Assessment of a remote sensing-based model for predicting malaria transmission risk in villages of Chiapas, Mexico. Am J Trop Med Hyg 1997;56 (1):99-106.

[50] Sharma VP, Dhiman RC, Ansari MA, Nagpal BN, Srivastava A, Manavalan P, Adiga S, Radhakrishnan K, Chandrasekhar MG. Study on the feasibility of delineating mosquitogenic conditions in and around Delhi using Indian Remote Sensing Satellite data. Indian J Malariol 1996;33 (3):107-25.

[51] Sharma VP, Srivastava A. Role of geographic information system in malaria control. Indian J Med Res 1997;106:198-204.

[52] Roberts DR, Paris JF, Manguin S, Harbach RE, Woodruff R, Rejmankova E, Polanco J, Wullschleger B, Legters LJ. Predictions of malaria vector distribution in Belize based on multispectral satellite data. Am J Trop Med Hyg 1996;54 (3):304-8.

[53] Roberts DR, Rodriguez MH. The environment, remote sensing, and malaria control. Ann N Y Acad Sci 1994;740:396-402.

[54] Rejmankova E, Roberts DR, Pawley A, Manguin S, Polanco J. Predictions of adult *Anopheles albimanus* densities in villages based on distances to remotely sensed larval habitats. Am J Trop Med Hyg 1995;53 (5):482-8.

[55] Lindsay SW, Parson L, Thomas CJ. Mapping the ranges and relative abundance of the two principal African malaria vectors, *Anopheles gambiae sensu stricto* and *An. arabiensis,* using climate data. Proc R Soc Lond B Biol Sci 1998;265 (1399):847-54.

[56] Da-peng L. Spatial prediction of malaria in Red River Basin, Yunnan, China using geographical information systems and remote sensing. Depertment of Infectious and Tropical Diseases. London: London School of Hygiene and Tropical Disease, 2000. pp. 214.

[57] Connor SJ, Thomson MC, Flasse SP, Perryman AH. Environmental information systems in malaria risk mapping and epidemic forecasting. Disasters 1998;22 (1):39-56.

[58] Connor SJ, Thomson MC, Molyneux DH. Forecasting and prevention of epidemic malaria: new perspectives on an old problem. Parassitologia 1999;41 (1-3):439-48.

[59] Thomas CJ, Lindsay SW. Local-scale variation in malaria infection amongst rural Gambian children estimated by satellite remote sensing. Trans R Soc Trop Med Hyg 2000;94 (2):159-63.

[60] Hay SI, Snow RW, Rogers DJ. Predicting malaria seasons in Kenya using multitemporal meteorological satellite sensor data. Trans R Soc Trop Med Hyg 1998;92 (1):12-20.

[61] Bailey TC, Gatrell AC. Interactive Spatial Data Analysis. Essex: Longman, 1995.

[62] Edrissian GH, Nateghpoor M, Afshar A, Sayedzadeh A, Mohsseni G, Satvat M, Emadi A. Monitoring the Response of *Plasmodium falciparum* and *P. vivax* to Antimalarial Drugs in the Malarious Areas in South-East Iran. Arch Iranian Med 1999;2 (2):2-7.

[63] Sardarizadeh S. Malaria in the world, in the Eastern Mediterranean region and in Iran. Arch Iranian Med 1999;2 (4):32-9.

[64] Edrissian GH. Status of the response of *Plasmodium falciparum* to chloroquine and mefloquine in Iran. Trop Geogr Med 1989;41 (4):297-303.

[65] Edrissian GH, Afshar A, Kanani A, Satvat MT, Mohsseni G, Nasseri Nejad K, Emadi AM, Ghorbani M. The response of *Plasmodium falciparum* to chloroquine and mefloquine in Bandar-Abbas and Minab areas, Hormozgan Province, southern Iran. J Trop Med Hyg 1989;92 (2):75-9.

[66] Edrissian GH, Afshar A, Sayedzadeh A, Mohsseni G, Satvat MT. Assessment of the response in vivo and in vitro of *Plasmodium falciparum* to sulphadoxine-pyrimethamine in the malarious areas of Iran. J Trop Med Hyg 1993;96 (4):237-40.

[67] WHO. Anti-malaria drug policies: Data requirements, treatment of uncomplicated malaria and management of malaria in pregrancy. World Health Organization/MAL/94, 1994. pp. 1070.

[68] WHO. Drug used in parasitic diseases: malaria. In: World Health Organization model prescribing information, 1995. pp. 24-73.

[69] Katsouyanni K, Schwartz J, Spix C, Touloumi G, Zmirou D, Zanobetti A, Wojtyniak B, Vonk JM, Tobias A, Ponka A, Medina S, Bacharova L, Anderson HR. Short term effects of air pollution on health: a European approach using epidemiologic time series data: the APHEA protocol. J Epidemiol Community Health 1996;50 Suppl 1:S12-8.

[70] Schwartz J, Spix C, Touloumi G, Bacharova L, Barumamdzadeh T, le Tertre A, Piekarksi T, Ponce de Leon A, Ponka A, Rossi G, Saez M, Schouten JP. Methodological issues in studies of air pollution and daily counts of deaths or hospital admissions. J Epidemiol Community Health 1996;50 Suppl 1:S3-11.

[71] Dominici F, Samet JM, Zegar SL. Combining evidence on air pollution and daily mortality from the 20 largest US cities: a hierarchical modelling strategy. J Roy Stat Soc 2000;163:263-302.

[72] Zakeri S, Najafabadi S, Zare A, Djadid N. Detection of malaria parasites by nested PCR in south-eastern, Iran: Evidence of highly mixed infections in Chahbahar district. Malar J 2002;1 (1):2.

[73] Singh B, Cox-Singh J, Miller AO, Abdullah MS, Snounou G, Rahman HA. Detection of malaria in Malaysia by nested polymerase chain reaction amplification of dried blood spots on filter papers. Trans R Soc Trop Med Hyg 1996;90 (5):519-21.

[74] Khoo A, Furuta T, Abdullah NR, Bah NA, Kojima S, Wah MJ. Nested polymerase chain reaction for detection of *Plasmodium falciparum* infection in Malaysia. Trans R Soc Trop Med Hyg 1996;90 (1):40-1.

[75] Hanscheid T, Grobusch MP. How useful is PCR in the diagnosis of malaria? Trends Parasitol 2002;18 (9):395-8.

[76] Barker RH, Jr., Banchongaksorn T, Courval JM, Suwonkerd W, Rimwungtragoon K, Wirth DF. *Plasmodium falciparum* and *P. vivax:* factors affecting sensitivity and specificity of PCR-based diagnosis of malaria. Exp Parasitol 1994;79 (1):41-9.

[77] Roper C, Elhassan IM, Hviid L, Giha H, Richardson W, Babiker H, Satti GM, Theander TG, Arnot DE. Detection of very low level *Plasmodium falciparum* infections using the nested polymerase chain reaction and a reassessment of the epidemiology of unstable malaria in Sudan. Am J Trop Med Hyg 1996;54 (4):325-31.

[78] Urdaneta L, Guevara P, Ramirez JL. Evaluation of DNA recombinant methodologies for the diagnosis of *Plasmodium falciparum* and their comparison with the microscopy assay. Mem Inst Oswaldo Cruz 1998;93 (5):639-46.

[79] Walsh PS, Metzger DA, Higuchi R. Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. Biotechniques 1991;10 (4):506-13.

[80] Boisier P, Jambou R, Raharimalala L, Roux J. Relationship between parasite density and fever risk in a community exposed to a low level of malaria transmission in Madagascar highlands. Am J Trop Med Hyg 2002;67 (2):137-40.

[81] Lyimo EO, Msuya FH, Rwegoshora RT, Nicholson EA, Mnzava AE, Lines JD, Curtis CF. Trial of pyrethroid impregnated bednets in an area of Tanzania holoendemic for malaria. Part 3. Effects on the prevalence of malaria parasitaemia and fever. Acta Trop 1991;49 (3):157-63.

[82] Lepers JP, Andriamagatiana-Rason MD, Coulanges P. [The treatment of *Plasmodium falciparum* malaria in the Malagasy Highland Plateaux. Development of fever and parasite density. Clinical aspects]. Arch Inst Pasteur Madagascar 1989;56 (1):155-9.

[83] Beadle C, McElroy PD, Oster CN, Beier JC, Oloo AJ, Onyango FK, Chumo DK, Bales JD, Sherwood JA, Hoffman SL. Impact of transmission intensity and age on *Plasmodium falciparum* density and associated fever: implications for malaria vaccine trial design. J Infect Dis 1995;172 (4):1047-54.

[84] Sexton JD, Deloron P, Bugilimfura L, Ntilivamunda A, Neill M. Parasitologic and clinical efficacy of 25 and 50 mg/kg of chloroquine for treatment of *Plasmodium falciparum* malaria in Rwandan children. Am J Trop Med Hyg 1988;38 (2):237-43.

[85] Plowe CV, Doumbo OK, Djimde A, Kayentao K, Diourte Y, Doumbo SN, Coulibaly D, Thera M, Wellems TE, Diallo DA. Chloroquine treatment of uncomplicated *Plasmodium falciparum* malaria in Mali: parasitologic resistance versus therapeutic efficacy. Am J Trop Med Hyg 2001;64 (5-6):242-6.

[86] Guthmann JP, Kasparian S, Phetsouvanh R, Nathan N, Garcia M, Phompida S, Brockman A, Gastellu M, Legros D. The efficacy of chloroquine for the treatment of acute, uncomplicated, *Plasmodium falciparum* malaria in Laos. Ann Trop Med Parasitol 2002;96 (6):553-7.

[87] Ehrhardt S, Mockenhaupt FP, Agana-Nsiire P, Mathieu A, Anemana SD, Stark K, Otchwemah RN, Bienzle U. Efficacy of chloroquine in the treatment of uncomplicated, *Plasmodium falciparum* malaria in northern Ghana. Ann Trop Med Parasitol 2002;96 (3):239-47.

[88] Dugelay F, Adehossi E, Adamou S, Ousmane I, Parzy D, Delmont J, Parola P. Efficacy of chloroquine in the treatment of uncomplicated, *Plasmodium falciparum* malaria in Niamey, Niger, in 2001. Ann Trop Med Parasitol 2003;97 (1):83-6.

[89] Jones TR, Baird JK, Basri H, Purnomo, Danudirgo EW. Prevalence of malaria in native and transmigrant populations. Effects of age and history of exposure. Trop Geogr Med 1991;43 (1-2):1-6.

[90] Voller A, Green DI, Richards WH. Cross immunity studies with East and West African strains of *Plasmodium falciparum* in owl monkeys (Aotus trivirgatus). J Trop Med Hyg 1973;76 (6):135-9.

[91] Weidanz WP, Long CA. The role of T cells in immunity to malaria. Prog Allergy 1988;41:215-52.

[92] Blomberg MT, Perlmann P. Malaria immunity: an overview with emphasis on T cell function. In: Good MF, Saul AJ, editors. Molecular immunological considerations in malaria vaccine development. Florida: CRC press, 1994. pp. 4-46.

[93] Smith T, Charlwood JD, Kihonda J, Mwankusye S, Billingsley P, Meuwissen J, Lyimo E, Takken W, Teuscher T, Tanner M. Absence of seasonal variation in malaria parasitaemia in an area of intense seasonal transmission. Acta Trop 1993;54 (1):55-72.

[94] Smith T, Charlwood JD, Kitua AY, Masanja H, Mwankusye S, Alonso PL, Tanner M. Relationships of malaria morbidity with exposure to *Plasmodium falciparum* in young children in a highly endemic area. Am J Trop Med Hyg 1998;59 (2):252-7.

[95] Boulos M, Amato Neto V, Dutra AP, Di Santi SM, Shiroma M. Frequency of malaria relapse due to *Plasmodium vivax* in a non-endemic region (Sao Paulo, Brazil). Rev Inst Med Trop Sao Paulo 1991;33 (2):143-6.

[96] Dua VK, Sharma VP. *Plasmodium vivax* relapses after 5 days of primaquine treatment, in some industrial complexes of India. Ann Trop Med Parasitol 2001;95 (7):655-9.

[97] Fernandopulle BM, Weeraratne CL, Weerasuriya K, Karunaweera ND. Efficacy of a five-day course of primaquine in preventing relapses in *Plasmodium vivax* malaria--a pilot study. Ceylon Med J 2003;48 (1):32.

[98] Gogtay NJ, Desai S, Kadam VS, Kamtekar KD, Dalvi SS, Kshirsagar NA. Relapse pattern of *Plasmodium vivax* in Mumbai: a study of 283 cases of *vivax* malaria. J Assoc Physicians India 2000;48 (11):1085-6.

[99] Prasad RN, Virk KJ, Sharma VP. Relapse/reinfection patterns of *Plasmodium vivax* infection: a four year study. Southeast Asian J Trop Med Public Health 1991;22 (4):499-503.

[100] Rowland M, Durrani N. Randomized controlled trials of 5- and 14-days primaquine therapy against relapses of *vivax* malaria in an Afghan refugee settlement in Pakistan. Trans R Soc Trop Med Hyg 1999;93 (6):641-3.

[101] Roy RG, Chakrapani KP, Dhinagaran D, Sitaraman NL, Ghosh RB. Efficacy of 5-day radical treatment of *P. vivax* infection in Tamil Nadu. Indian J Med Res 1977;65 (5):652-6.

[102] Sharma RC, Gautam AS, Orlov V, Sharma VP. Relapse pattern of *Plasmodium vivax* in Kheda district, Gujarat. Indian J Malariol 1990;27 (2):95-9.

[103] Srivastava HC, Sharma SK, Bhatt RM, Sharma VP. Studies on *Plasmodium vivax* relapse pattern in Kheda district, Gujarat. Indian J Malariol 1996;33 (4):173-9.

[104] Mason J. Patterns of *Plasmodium vivax* recurrence in a high-incidence coastal area of El Salvador, C. A. Am J Trop Med Hyg 1975;24 (4):581-5.

[105] Adak T, Sharma VP, Orlov VS. Studies on the *Plasmodium vivax* relapse pattern in Delhi, India. Am J Trop Med Hyg 1998;59 (1):175-9.

[106] Leslie T, Rab MA, Ahmadzai H, Durrani N, Fayaz M, Kolaczinski J, Rowland M. Compliance with 14-day primaquine therapy for radical cure of *vivax* malaria -a randomized placebo-controlled trial comparing unsupervised with supervised treatment. Trans R Soc Trop Med Hyg 2004;98 (3):168-73.

[107] Yadav RS, Ghosh SK. Radical curative efficacy of five-day regimen of primaquine for treatment of *Plasmodium vivax* malaria in India. J Parasitol 2002;88 (5):1042-4.

[108] Durante Mangoni E, Severini C, Menegon M, Romi R, Ruggiero G, Majori G. Case report: An unusual late relapse of *Plasmodium vivax* malaria. Am J Trop Med Hyg 2003;68 (2):159-60.

[109] WHO. Resistance of malaria parasites to drugs. Geneva: World Health Organization, Division of control of tropical dieases, 1995.

[110] WHO. Assessment of therapeutic efficacy of antimalarial drugs for uncomplicated *falciparum* malaria in areas with intence transmission. Geneva: World Health Organization, Division of control of tropical dieases, 1996.

[111] Fontanet AL, Walker AM. Predictors of treatment failure in multiple drug-resistant *falciparum* malaria: results from a 42-day follow-up of 224 patients in eastern Thailand. Am J Trop Med Hyg 1993;49 (4):465-72.

[112] Sumawinata IW, Bernadeta, Leksana B, Sutamihardja A, Purnomo, Subianto B, Sekartuti, Fryauff DJ, Baird JK. Very high risk of therapeutic failure with chloroquine for uncomplicated *Plasmodium falciparum* and *P. vivax* malaria in Indonesian Papua. Am J Trop Med Hyg 2003;68 (4):416-20.

[113] Watt G, Loesuttiviboon L, Long GW. Prospective comparison of methods for the early prediction of treatment failure in patients with *falciparum* malaria. Clin Infect Dis 1995;21 (4):1026-8.

[114] Rowland M, Durrani N, Hewitt S, Sondorp E. Resistance of *falciparum* malaria to chloroquine and sulfadoxine-pyrimethamine in Afghan refugee settlements in western Pakistan: surveys by the general health services using a simplified in vivo test. Trop Med Int Health 1997;2 (11):1049-56.

[115] Edrissian GH, Shahabi S. Preliminary study of the response of *Plasmodium falciparum* to chloroquine in Sistan and Baluchestan province of Iran. Trans R Soc Trop Med Hyg 1985;79 (4):563-4.

[116] Edrissian GH, Ghorbani M, Afshar A, Kanani A, Satvat MT. In vitro response of *Plasmodium falciparum* to mefloquine in south-eastern Iran. Trans R Soc Trop Med Hyg 1987;81 (1):164-5.

[117] Eskandarian AA, Keshavarz H, Basco LK, Mahboudi F. Do mutations in *Plasmodium falciparum* dihydropteroate synthase and dihydrofolate reductase confer resistance to sulfadoxine-pyrimethamine in Iran? Trans R Soc Trop Med Hyg 2002;96 (1):96-8.

[118] Hamedi Y, Nateghpour M, Tan-ariya P, Tiensuwan M, Silachamroon U, Looareesuwan S. *Plasmodium vivax* malaria in Southeast Iran in 1999-2001: establishing the response to chloroquine in vitro and in vivo. Southeast Asian J Trop Med Public Health 2002;33 (3):512-8.

[119] Jafari S, Le Bras J, Asmar M, Durand R. Molecular survey of *Plasmodium falciparum* resistance in south-eastern Iran. Ann Trop Med Parasitol 2003;97 (2):119-24.

[120] Suroso T, Hamidi AN, Manouchehri AV. The activity of chloroquine against *Plasmodium falciparum* in Bandar Abbas, Southern Iran, 1976. Bull Soc Pathol Exot Filiales 1978;71 (2):164-71.

[121] Zakeri S, Gil JP,. Bereckzy S, Djadid ND, Bjorkman A. High prevalence of double *Plasmodium falciparum* dhfr mutations at codons 108 and 59 in the Sistan-Baluchistan province, Iran. J Infect Dis 2003;187 (11):1828-9.

[122] Duarte EC, Pang LW, Ribeiro LC, Fontes CJ. Association of subtherapeutic dosages of a standard drug regimen with failures in preventing relapses of *vivax* malaria. Am J Trop Med Hyg 2001;65 (5):471-6.

[123] Gill CA. The seasonal periodicity of malaria and the mechanism of the epidemic wave. London: Churchill, 1938.

[124] Macdonald G. The analysis of malaria epidemics. Trop Dis Bull 1953;50:871-89.

[125] Howard SC, Donnell CA, MS C. Methods for estimation of associations between multiple species parasite infections. Parasitology 2001;122:223-51.

[126] McKenzie FE, Bossert WH. Mixed-species *Plasmodium* infections of humans. J Parasitol 1997;83 (4):593-600.

[127] McKenzie FE, Bossert WH. Mixed-species *Plasmodium* infections of *Anopheles* (Diptera:Culicidae) [published erratum appears in J Med Entomol 1997 Sep;34(5):ii]. J Med Entomol 1997;34 (4):417-25.

[128] McKenzie FE, Bossert WH. Multispecies *Plasmodium* infections of humans. J Parasitol 1999;85 (1):12-8.

[129] Molineaux L, Gramiccia G. The study design and study area. In: The Garki Project, research on the epidemiology and control of malaria in the Sudan Savanna of west Africa. Geneva: World Health Organization, 1980. pp. 109-72.

[130] Cohen JE. Heterologous immunity in human malaria. Q Rev Biol 1973;48 (3):467-89.

[131] Smith T, Genton B, Baea K, Gibson N, Narara A, Alpers MP. Prospective risk of morbidity in relation to malaria infection in an area of high endemicity of multiple species of *Plasmodium*. Am J Trop Med Hyg 2001;64 (5-6):262-7.

[132] Pinto J, Sousa CA, Gil V, Goncalves L, Lopes D, do Rosario VE, Charlwood JD. Mixed-species malaria infections in the human population of Sao Tome island, west Africa. Trans R Soc Trop Med Hyg 2000;94 (3):256-7.

[133] Boyd MF, F kS. Veneral *vivax* activity in persons simulataneously inoculated with *Plasmodium vivax* and *Plaspodium falciparum*. Am J Trop Med Hyg 1938;18:505-14.

[134] Bruce MC, Donnelly CA, Alpers MP, Galinski MR, Barnwell JW, Walliker D, Day KP. Cross-species interactions between malaria parasites in humans. Science 2000;287 (5454):845-8.

[135] Looareesuwan S, White NJ, Chittamas S, Bunnag D, Harinasuta T. High rate of *Plasmodium vivax* relapse following treatment of *falciparum* malaria in Thailand. Lancet 1987;2 (8567):1052-5.

[136] Maitland K, Williams TN, Bennett S, Newbold CI, Peto TE, Viji J, Timothy R, Clegg JB, Weatherall DJ, Bowden DK. The interaction between *Plasmodium falciparum* and *P. vivax* in children on Espiritu Santo island, Vanuatu. Trans R Soc Trop Med Hyg 1996;90 (6):614-20.

[137] Mason DP, McKenzie FE, Bossert WH. The blood-stage dynamics of mixed *Plasmodium* malariae-*Plasmodium falciparum* infections. J Theor Biol 1999;198 (4):549-66.

[138] Molineaux L, Gramiccia G. Parasitology. In: The Garki Project, research on the epidemiology and control of malaria in the Sudan Savanna of west Africa. Geneva: World Health Organization, 1980. pp. 109-72.

[139] Molineaux L, Storey J, Cohen JE, Thomas A. A longitudinal study of human malaria in the West African Savanna in the absence of control measures: relationships between different *Plasmodium* species, in particular *P. falciparum* and *P. malariae.* Am J Trop Med Hyg 1980;29 (5):725-37.

[140] Richie TL. Interactions between malaria parasites infecting the same vertebrate host. Parasitology 1988;96 (Pt 3):607-39.

[141] Fox E, Strickland GT. The interrelationship of *Plasmodium falciparum* and *P. vivax* in the Punjab. Trans R Soc Trop Med Hyg 1989;83 (4):471-3.

[142] Farmer JN, Breitenbach RP. Cross immunity between *Plasmodium relictum* and *Plasmodium lophurae* in chickens. Poult Sci 1969;48 (3):785-91.

[143] Thompson SG. Why sources of heterogeneity in meta-analysis should be investigated. Br Med J 1994;309 (6965):1351-5.

[144] Thompson SG, Sharp SJ. Explaining heterogeneity in meta-analysis: a comparison of methods. Stat Med 1999;18 (20):2693-708.

[145] Hardy RJ, Thompson SG. Detecting and describing heterogeneity in meta-analysis. Stat Med 1998;17 (8):841-56.

[146] Cross JH, Clarke MD, Durfee PT, Irving GS, Taylor J, Partono F, Joesoef A, Hudojo, Oemijati S. Parasitology survey and seroepidemiology of amoebiasis in South Kalimantan (Borneo), Indonesia. Southeast Asian J Trop Med Public Health 1975;6 (1):52-60.

[147] Giboda M, Pholsena K, Hongvanthong B, Gutvirth J, Rubik I. Malariometric survey in Keoudom district, Laos: sensitivity of *Plasmodium falciparum* to antimalarials and automedication with chloroquine. Southeast Asian J Trop Med Public Health 1992;23 (3):383-8.

[148] Maffi M, Hutapea AM, Supardi P, Hadi S. On an unorthodox malariometric survey around Bokondini (Jayawijaya), Irian Jaya, Indonesia. Trans R Soc Trop Med Hyg 1975;36 (4):233-54.

[149] Anthony RL, Bangs MJ, Hamzah N, Basri H, Purnomo, Subianto B. Heightened transmission of stable malaria in an isolated population in the highlands of Irian Jaya, Indonesia. Am J Trop Med Hyg 1992;47 (3):346-56.

[150] Rajagopal R. Studies on persistent transmission of malaria in Burnihat, Meghalaya. J Commun Dis 1976;8 (4):235-45.

[151] Syafruddin, Kamimura K, Hasegawa H, Toma T, Miyagi I, Kawamoto F, Nainggolan IJ, Tumewu-Wagey M, Mandagi-Waworuntu H, Kapojos FX, Runtuwene J. Epidemiological study of malaria in north Sulawesi, Indonesia by fluorescence and Giemsa staining. Jap J Med Sci Biolo 1992;45 (4):175-84.

[152] Cross JH, Clarke MD, Cole WC, Lien JC, Partono F, Joesoef A, Kosin EH. Parasitology survey in northern Sumatra, Indonesia. J Trop Med Hyg 1976;79 (6):123-31.

[153] Adak T, Batra CP, Mittal PK, Sharma VP. Epidemiological study of malaria outbreak in a hotel construction site of Delhi. Indian J Malariol 1994;31 (3):126-31.

[154] Smrkovski LL, Escamilla J, Wooster MT, Rivera DG. A preliminary survey of malaria in Occidental Mindoro, Philippines. Southeast Asian J Trop Med Public Health 1982;13 (2):181-5.

[155] Dutta P, Bhattacharyya DR, Khan SA, Sharma CK, Goswami BK. Some observations of malaria in Boko PHC of Kamrup District, Assam. J Commun Dis 1994;26 (1):52-5.

[156] Cattani J, Taufa T, Anderson W, Lourie J. Malaria and filariasis in the Ok Tedi Region of the Star Mountains, Papua New Guinea. P N G Med J 1983;26 (2):122-6.

[157] Rafi S, Memon MA, Rao MH, Billoo AG. A change of *Plasmodium* species infecting children in Karachi over the last decade. J Pak Med Assoc 1994;44 (7):162-4.

[158] Hii J, Kan S, Pereira M, et al. Bancroftian filariasis and malaria in island and hinterland populations in Sabah, Malaysia. Trop Geographical Med 1985;37 (2):93-101.

[159] Das S, Malakar P, Saha GK, Dasgupta B, Hati AK. An epidemiological and entomological survey on malaria in an endemic area of Jalpaiguri district, North Bengal. J Bengal Natural History Society 1994;13 (1):75-6.

[160] Mizushima Y, Kato H, Ohmae H, Tanaka T, Bobogare A, Ishii A. Prevalence of malaria and its relationship to anemia, blood glucose levels, and serum somatomedin c (IGF-1) levels in the Solomon Islands. Acta Trop 1994;58 (3/4):207-20.

[161] De Arruda M, Carvalho MB, Nussenzweig RS, et al. Potential vectors of malaria and their different susceptibility to *Plasmodium falciparum* and *Plasmodium vivax* in northern Brazil identified by immunoassay. Am J Trop Med Hyg 1986;35 (5):873-81.

[162] Gautret P, Barreto M, Mendez F, Zorrilla G, Carrasquilla G. High prevalence of malaria in a village of the Colombian Pacific coast. Mem Inst Oswaldo Cruz 1995;90 (5):559-60.

[163] Cattani JA, Tulloch JL, Vrbova H, et al. The epidemiology of malaria in a population surrounding Madang, Papua New Guinea. Am J Trop Med Hyg 1986;35 (1):3-15.

[164] Ghosh SK, Yadav RS. Naturally acquired concomitant infections of bancroftian filariasis and human *plasmodia* in Orissa. Indian J Malariol 1995;32 (1):32-6.

[165] Verma BL, Srivastava RN. Quantitative assessment of malaria morbidity based on longitudinal data in 10 Indian villages. J Trop Med Hyg 1986;89 (2):57-60.

[166] Dietze R, Perkins M, Bóulos M, Luz F, Reller B, Corey GR. The diagnosis of *Plasmodium falciparum* infection using a new antigen detection system. Am J Trop Med Hyg 1995;52 (1):45-9.

[167] Beljaev AE, Brohult JA, Sharma GK, Samantaray KC. Studies on the detection of malaria at primary health centres. Part III. Parasitological profile of population surveyed for malaria through passive case detection. Indian J Malariol 1987;24 (2):97-106.

[168] Dutta P, Mahanta J. Incrimination of *Anopheles minimus* as a vector of malaria in Karbi Anglong district of Assam. Indian J Malariol 1995;32 (3):129-31.

[169] Strickland GT, Zafar Latif A, Fox E, Khaliq AA, Chowdhry MA. Endemic malaria in four villages of the Pakistani province of Punjab. Trans R Soc Trop Med Hyg 1987;81 (1):36-41.

[170] Uchida JY, Kasahara T, Bobogare A, Saefafia S, Kere N, Kawabata M, Ohta N, Ishii A. The prevalence of *falciparum* malaria in the Solomon Islands investigated by a filter paper disk-PCR method. Jap J Parasitol 1995;44 (2):119-27.

[171] Sherchand JB, Ohara H, Shrestha MP, Sherchand S. Association of G-6-PD deficiency sickle cell haemoglobin and blood groups with resistance to malaria infection. A survey on different castes in southern Nepal. Jap J Parasitol 1995;44 (5):396-403.

[172] Das NG, Baruah I, Kamal S, Sarkar PK, Das SC, Santhanam K. An epidemiological and entomological investigation on malaria outbreak at Tamulpur PHC, Assam. Indian J Malariol 1997;34 (3):164-70.

[173] Strickland GT, Fox E, Hadi H. Malaria and splenomegaly in the Punjab. Trans R Soc Trop Med Hyg 1988;82 (5):667-70.

[174] Hansmann Y, Staub-Schmidt T, Christmann D. Imported malaria in Strasbourg: an epidemiological, clinical, biological and therapeutic study. P N G Med J 1997;2 (10):941-52.

[175] Singh N, Sharma VP. Persistent malaria transmission in Kundam block, District Jabalpur (M.P.). Indian J Malariol 1989;26 (1):1-7.

[176] Belizario VY, Saul A, Bustos MDG, Lansang MA, Pasay CJ, Gatton M, Salazar NP. Field epidemiological studies on malaria in a low endemic area in the Philippines. Acta Trop 1997;63 (4):241-56.

[177] Dutta P, Bhattacharyya DR, Dutta LP. Incrimination of *Anopheles dirus* as a vector of malaria in Dibrugarh district, Assam. Indian J Malariol 1989;26 (3):149-52.

[178] Itokawa H, Takai R, Ishii A, Panjaitan W. Age specific asexual parasite and gametocyte density in highly endemic malaria in North Sumatra, Indonesia. Jap J Trop Med Hyg 1989;17 (4):303-10.

[179] Seboxa T, Snow RW. Epidemiological features of severe paediatric malaria in north western Ethiopia. East Africa Med J 1997;74 (12):780-3.

[180] Ghosh SK, Kumar A, Chand SK, Choudhury DS. A preliminary malaria survey in Bisra PHC, District Sundergarh, Orissa. Indian J Malariol 1989;26 (3):167-70.

[181] Carney WP, Van Peenen PF, See R, Hagelstein E, Lima B. Parasites of man in remote areas of Central and South Sulawesi, Indonesia. Southeast Asian J Trop Med Public Health 1977;8 (3):380-9.

[182] Graves PM, Eida S, Lagog M. Malaria in adult outpatients at Goroka Hospital during 1986. P N G Med J 1989;32 (3):189-93.

[183] Mandal B, Mitra NK, Mukhopadhyay AK, Mukherjee H, Hati AK. Emerging *Plasmodium falciparum* in an endemic area in Calcutta. J Ind Med Assoc 1998;96 (11):328-9.

[184] Joshi H, Malhotra MS, Raghavendra K, Subbarao SK, Sharma VP. Genetic studies among Buksa tribals. J Para Dis 1998;22 (2):136-9.

[185] Das LK, Mohapatra SS, Jambulingam P, Gunasekaran K, Pani SP, Das PK. Malaria and other common ailments among upper Bonda tribals in Koraput district, Orissa. Indian J Med Res 1989:89334-9.

[186] Camargo LMA, Noronha E, Salcedo JMV, Dutra AP, Krieger H, Silva LHPd, Camargo EP. The epidemiology of malaria in Rondonia (Western Amazon region, Brazil): study of a riverine population. Acta Trop 1999;72 (1):1-11.

[187] Dutta P, Bhattacharyya DR. Malaria survey in some parts of Namsang Circle of Tirap District, Arunachal Pradesh. J Commun Dis 1990;22 (2):92-7.

[188] Singh N, Mishra SS, Singh MP, Sharma VP. Seasonality of *Plasmodium vivax* and *P. falciparum* in tribal villages in central India (1987-1995). Curr Sci 2000;94 (2):101-12.

[189] Rajagopalan PK, Das PK, Pani SP. Parasitological aspects of malaria persistence in Koraput district, Orissa, India. Indian J Med Res 1990;91 (January):44-51.

[190] Hozhabri S, Akhtar S, Rahbar MH, Luby SP. Prevalence of *plasmodium* slide positivity among the children treated for malaria, Jhangara, Sindh. J Pak Med Assoc 2000;50 (12):401-5.

[191] Nosten F, Ter Kuile F, Maelankirri L, Decludt B, White NJ. Malaria during pregnancy in an area of unstable endemicity. Trans R Soc Trop Med Hyg 1991;85 (4):424-9.

[192] Mehlotra RK, Lorry K, Kastens W, Miller SM, Alpers MP, Bockarie M, Kazura JW, Zimmerman PA. Random distribution of mixed species malaria infections in Papua New Guinea. Am J Trop Med Hyg 2000;62 (2):225-31.

[193] Subramanian S, Manoharan A, Sahu S, Jambulingam P, Govardhini P, Mohapatra SSS, Das PK. Living conditions and occurrence of malaria in a rural community. Indian J Malariol 1991;28 (1):29-37.

[194] Srivastava HC, Yadav RS. Malaria outbreak in a tribal area of Gujarat state, India. Southeast Asian J Trop Med Public Health 2000;31 (2):219-24.

[195] Moitinho MdLR, Sobrinho AN, Casavechia MTG, Silva Filho VL, Lima EMd, Souza MId, Leal EF. A study of 338 patients with suspected malaria in Maringa-Parana, Brazil. Rev UNIMAR 1991;13 (1):31-6.

[196] Pinto J, Sousa CA, Gil V, Goncalves L, Lopes D, Rosario VEd, Charlwood JD. Mixed-species malaria infections in the human population of Sao Tome island, West Africa. Trans R Soc Trop Med Hyg 2000;94 (3):256-7.

[197] Gordon DM, Davis DR, Lee M. Significance of circumsporozoite-specific antibody in the natural transmission of *Plasmodium falciparum*, *Plasmodium vivax*, and *Plasmodium malariae* in an aboriginal (Orang Asli) population of central peninsular Malaysia. Am J Trop Med Hyg 1991;45 (1):49-56.

[198] Roper MH, Carrion Torres RS, Cava Goicochea CG, Andersen EM, Aramburu Guarda JS, Calampa C, Hightower AW, Magill AJ. The epidemiology of malaria in an epidemic area of the Peruvian Amazon. Am J Trop Med Hyg 2000;62 (2):247-56.

[199] Dutta P, Bhattacharyya DR, Dutta LP. Epidemiological observations on malaria in some parts of Tengkhat PHC, Dibrugarh district, Assam. Indian J Malariol 1991;28 (2):121-8.

[200] Singh N, Mehra RK, Srivastava N. Malaria during pregnancy and infancy, in an area of intense malaria transmission in central India. Ann Trop Med Parasitol 2001;95 (1):19-29.

[201] Molineaux L, Gramiccia G. The mathematical model of transmission. In: The Garki Project, research on the epidemiology and control of malaria in the Sudan Savanna of west Africa. Geneva: World Health Organization, 1980. pp. 261-88.

[202] Neuhaus JM. Statistical methods for longitudinal and clustered designs with binary responses. Stat Methods Med Res 1992;1 (3):249-73.

[203] Bekessy A, Molineaux L, Storey J. Estimation of incidence and recovery rates of *Plasmodium falciparum* parasitaemia from longitudinal data. Bull World Health Organ 1976;54 (6):685-93.

[204] Goldstein H. The basic linear multilevel model and its estimation. In: Multi-level statistical models. London: Arnold, 1995.

[205] Boyd MF. Malariology. Philadelphia and London: Saunders, 1949.

[206] Kitua AY, Smith T, Alonso PL, Masanja H, Urassa H, Menendez C, Kimario J, Tanner M. *Plasmodium falciparum* malaria in the first year of life in an area of intense and perennial transmission. Trop Med Int Health 1996;1 (4):475-84.

[207] Sehgal VM, Siddjiqui WA, Alpers MP. A seroepidemiological study to evaluate the role of passive maternal immunity to malaria in infants. Trans R Soc Trop Med Hyg 1989;83 Suppl:105-6.

[208] Thomas TC. Biting activity of *Anopheles gambiae*. Br Med J 1951;4744:1402.

[209] Muirhead-Thomson RC. The distribution of *anopheline* mosquito bites among different age groups; a new factor in malaria epidemiology. Br Med J 1951;4715:1114-7.

[210] Lindsay SW, Adiamah JH, Miller JE, Pleass RJ, Armstrong JR. Variation in attractiveness of human subjects to malaria mosquitoes (Diptera: Culicidae) in The Gambia. J Med Entomol 1993;30 (2):368-73.

[211] Molineaux L, Gramiccia G. Immunology. In: The Garki Project, research on the epidemiology and control of malaria in the Sudan Savanna of west Africa. Geneva: World Health Organization, 1980. pp. 173-212.

[212] Hafalla JC, Morrot A, Sano G, Milon G, Lafaille JJ, Zavala F. Early self-regulatory mechanisms control the magnitude of CD8+ T cell responses against liver stages of murine malaria. J Immunol 2003;171 (2):964-70.

[213] Reece WH, Pinder M, Gothard PK, Milligan P, Bojang K, Doherty T, Plebanski M, Akinwunmi P, Everaere S, Watkins KR, Voss G, Tornieporth N, Alloueche A, Greenwood BM, Kester KE, McAdam KP, Cohen J, Hill AV. A CD4(+) T-cell immune response to a conserved epitope in the circumsporozoite protein correlates with protection from natural *Plasmodium falciparum* infection and disease. Nat Med 2004;10 (4):406-10.

[214] Anderson RJ, Hannan CM, Gilbert SC, Laidlaw SM, Sheu EG, Korten S, Sinden R, Butcher GA, Skinner MA, Hill AV. Enhanced CD8+ T cell immune responses and protection elicited against *Plasmodium berghei* malaria by prime boost immunization regimens using a novel attenuated fowlpox virus. J Immunol 2004;172 (5):3094-100.

[215] Troye-Blomberg M, Perlmann P. T cell functions in *Plasmodium falciparum* and other malarias. Prog Allergy 1988;41:253-87.

[216] Ferrante A, Kumaratilake LM, rathjen DA. The role of cytokine-activated phagocytic cell immunity to malaria. In: Good MF, Saul AJ, editors. Molecular immunological considerations in malaria vaccine development. Florida: CRC press, 1994. pp. 48-86.

[217] Celada A, Cruchaud A, Perrin LH. Independence of complement on in vitro immune phagocytosis of *Plasmodium falciparum* parasitised erythrocytes by human monocytes and polymorphonuclear leukocytes. Int Arch Allergy Appl Immunol 1984;73 (4):363-6.

[218] Kumaratilake LM, Ferrante A, Rzepczyk CM. Tumor necrosis factor enhances neutrophil-mediated killing of *Plasmodium falciparum.* Infect Immun 1990;58 (3):788-93.

[219] Garraud O, Mahanty S, Perraut R. Malaria-specific antibody subclasses in immune individuals: a key source of information for vaccine design. Trends Immunol 2003;24 (1):30-5.

[220] Singer LM, Mirel LB, Kuile FO, Branch OH, Vulule JM, Kolczak MS, Hawley WA, Kariuki SK, Kaslow DC, Lanar DE, Lal AA. The Effects of Varying Exposure to Malaria Transmission on Development of Antimalarial Antibody Responses in Preschool Children. XVI. Asembo Bay Cohort Project. J Infect Dis 2003;187 (11):1756-64.

[221] Baird JK. Age-dependent characteristics of protection v. susceptibility to *Plasmodium falciparum.* Ann Trop Med Parasitol 1998;92 (4):367-90.

[222] Baird JK, Jones TR, Danudirgo EW, Annis BA, Bangs MJ, Basri H, Purnomo, Masbar S. Age-dependent acquired protection against *Plasmodium falciparum* in people having two years exposure to hyperendemic malaria. Am J Trop Med Hyg 1991;45 (1):65-76.

[223] Baird JK, Purnomo, Basri H, Bangs MJ, Andersen EM, Jones TR, Masbar S, Harjosuwarno S, Subianto B, Arbani PR. Age-specific prevalence of *Plasmodium falciparum* among six populations with limited histories of exposure to endemic malaria. Am J Trop Med Hyg 1993;49 (6):707-19.

[224] Baird JK, Masbar S, Basri H, Tirtokusumo S, Subianto B, Hoffman SL. Age-dependent susceptibility to severe disease with primary exposure to *Plasmodium falciparum.* J Infect Dis 1998;178 (2):592-5.

[225] Troye-Blomberg M, Olerup O, Larsson A, Sjoberg K, Perlmann H, Riley E, Lepers JP, Perlmann P. Failure to detect MHC class II associations of the human immune response induced by repeated malaria infections to the *Plasmodium falciparum* antigen Pf155/RESA. Int Immunol 1991;3 (10):1043-51.

[226] Sinigaglia F, Guttinger M, Romagnoli P, Takacs B. Malaria antigens and MHC restriction. Immunol Lett 1990;25 (1-3):265-70.

[227] Widmann C, Romero P, Maryanski JL, Corradin G, Valmori D. T helper epitopes enhance the cytotoxic response of mice immunized with MHC class I-restricted malaria peptides. J Immunol Methods 1992;155 (1):95-9.

[228] Riley EM. The role of MHC- and non-MHC-associated genes in determining the human immune response to malaria antigens. Parasitology 1996;112 Suppl:S39-51.

[229] Rui-Mei L, Kara AU, Sinniah R. Upregulation of major histocompatibility complex (MHC) antigen in nephritis associated with murine malaria infection. J Pathol 1998;185 (2):212-8.

[230] Luyendyk J, Olivas OR, Ginger LA, Avery AC. Antigen-presenting cell function during *Plasmodium yoelii* infection. Infect Immun 2002;70 (6):2941-9.

[231] Cifuentes G, Patarroyo ME, Urquiza M, Ramirez LE, Reyes C, Rodriguez R. Distorting Malaria Peptide Backbone Structure to Enable Fitting into MHC Class II Molecules Renders Modified Peptides Immunogenic and Protective. J Med Chem 2003;46 (11):2250-3.

[232] Andersen E, Jones TR, Purnomo, Masbar S, Wiady I, Tirtolusumo S, Bangs MJ, Charoenvit Y, Gunawan S, Hoffman SL. Assessment of age-dependent immunity to malaria in transmigrants. Am J Trop Med Hyg 1997;56 (6):647-9.