Estimating Malaria Incidence and Recovery Rates from Panel Surveys

BURTON SINGER*

AND

JOEL E. COHEN

The Rockefeller University, 1230 York Avenue, New York, New York 10021

Received 12 November 1979

ABSTRACT

We estimate and interpret two forms of age- and season-specific malaria infection and recovery rates. Our analysis is based on a time series of two-wave panel surveys from the first longitudinal field study that allows for direct estimation of these quantities—a WHO malaria survey in Garki, Kano State, Nigeria. We present a strategy for deciding whether two-state, two-wave panel data could have been generated by continuous-time Markov chains or certain mixtures of such chains. The central idea, applicable to longitudinal surveys generally, is to test estimated conditional probabilities for membership in the set of conditional probabilities that can be generated by specific classes of models (e.g. inhomogeneous Markov chains). Our method assumes *a priori* only that the observations arise by sampling some continuous-time stochastic process.

1. INTRODUCTION

The purpose of this paper is to estimate and interpret conversion (or incidence) and recovery rates in human populations exposed to the malarial parasite *Plasmodium falciparum*. Our estimates are based on a time series of two-wave panel surveys. These surveys are part of a longitudinal field study of malaria conducted in Garki, Kano State, Nigeria by the Government of the Federal Republic of Nigeria and the World Health Organization [21].

The basic objectives of the Garki project were to (1) study the epidemiology of malaria, (2) measure the impact of certain control measures, and (3) construct and test a mathematical model allowing simulation of the trans-

*Also with Columbia University, New York, NY 10027.

52 Vanderbilt Ave., New York, NY 10017. 0025-5564/80/040273+33$02.25

mission of malaria. Our primary concern is with (1). Objectives (2) and (3) are pursued elsewhere [22, 8, 10, 11, 21].

The two forms of conversion and recovery rates presented here are estimated from baseline data collected prior to any attempts to control malaria. Our investigation was motivated by an interesting paper by Bekessy et al. [4] on the estimation of one form of such rates; namely, the expected number of events (conversion or recovery) per unit time per person at risk of the event at a time t. We also estimate and interpret the expected numbers of conversions and recoveries per unit time per person in the survey population. These rates represent information complementary to that discussed by Bekessy et al. In particular, the rates per person in the survey population reflect the influence of the relative sizes of the reservoirs of positive and negative persons at the initial survey. A limiting form of these rates may be interpreted as steady-state epidemiological incidence rates.

In Sec. 2 we review the characteristics of malarial infections and describe the role of the Garki project in the historical context of field studies of malaria. We also indicate the importance of conversion and recovery rates for an understanding of malaria in human populations and for the assessment of the impact of interventions aimed at either control or eradication of the disease.

This work also contributes to the larger problem of developing concepts and models to analyze longitudinal data. For a prescribed event A, associated with a continuous-time counting process, natural formulations of the concepts "event rate at time s per individual in the total population" and "event rate at time s per individual at risk at time s" are, respectively,

$$r_A(s) = \lim_{t \downarrow s} \frac{E(\text{number of occurrences of } A \text{ per individual in } [s,t))}{t-s} \tag{1.1}$$

and

$$\rho_A(s) = \lim_{t \downarrow s} \frac{E(\text{number of occurrences of } A \text{ per individual in } [s,t) \mid \text{individual is at risk of } A \text{ at time } s)}{t-s} \tag{1.2}$$

In empirical applications, testing whether classes of counting processes describe the occurrence of events in a population and estimating (1.1) and (1.2) within such a class are best facilitated by observing, in full, N realizations of a counting process

$$\nu(t, w_i), \qquad 0 < t < T, \quad 1 \leqslant i \leqslant N. \tag{1.3}$$

For each individual i, $\nu(t, w_i)$ is the number of occurrences of event A by time t starting from zero occurrences at the initial time, N is the number of individuals in a survey, and $T \gg \inf(t : \nu(t, w) > 0)$ with high probability.

In many longitudinal surveys, collecting the exact timing of each occurrence of an event for each individual as in (1.3) is either impossible, economically infeasible, or both. Observations usually contain gaps and censoring relative to a continuously evolving process [27]. For example, in a multiwave panel survey, observations on characteristics of a population are taken in waves at discrete time points $0 = t_0 < t_1 < \cdots < t_n$, and several events in a counting process $\nu(t, w)$ could occur in time $\Delta = \min_i(t_{i+1} - t_i)$.

Section 3 describes the connection between hypothesis testing for two-wave panel data and the embedding problem for Markov chains. We present the first systematic treatment of sampling aspects of embeddability criteria. We study the variability of parameter estimators for two-state continuous-time homogeneous Markov chains based on two-wave panel data. These estimators are then used to estimate conversion and recovery rates within the class of Markov models, once the embeddability of the panel observations within this class of models is established. This is the prototype of a general strategy for testing whether proposed classes of stochastic processes could have generated given multi-wave panel data. Some unsolved problems of such a strategy and the related literature on inference for stochastic processes are discussed in Sec. 8.

In Sec. 4 we apply our methods to data studied previously [4]. The Markovian parameter estimates are used in Sec. 5 to estimate the event rates (1.1) and (1.2) within this class of models. In Sec. 6 we interpret these rates in the context of the Garki baseline surveys.

The measurement error that arises in parasitemia surveys from daily oscillations in parasite density in the blood of infected persons is modeled in Sec. 7. We show that this measurement error causes a downward bias in estimated equilibrium conversion and recovery rates. We illustrate some biases that can occur in the estimation of event rates due to model misspecification. Since the criterion for embedding two-wave panel data in continuous-time Markov chains is identical to the criterion for embedding such data in a wide class of nondegenerate mixtures of such chains, the estimation of conversion and recovery rates assuming a Markov-chain model could be severely biased if a nondegenerate mixture of such chains actually generated the data.

2. QUANTITATIVE STUDIES OF MALARIA

2.1. *NATURAL HISTORY OF INFECTIONS*

Malaria is an acute and chronic infection caused by protozoa of the genus *Plasmodium*. Malarial infections are characterized by fever,

splenomegaly (enlargement of the spleen), anemia, and, in children, frequently fatal complications. Four species of *Plasmodium* occur naturally in man: *P. falciparum*, *P. malariae*, *P. ovale*, and *P. vivax*. The first three of these species are present in the Garki district of northern Nigeria. This paper concerns only infections due to the most prevalent species, *P. falciparum*. To clarify how parasitological measurements—the raw material for the estimation of conversion and recovery rates—relate to the infection history of an individual, we describe briefly the life cycle of *P. falciparum*. For details and scientific history, see [26] and [25]. For clinical aspects, see [30].

The life cycle of malarial parasites consists of a sexual phase in certain anopheles mosquitoes—notably *Anopheles gambiae* in Garki—and an asexual phase with multiplication in man. The anopheline phase of the cycle begins when the mosquito ingests human blood containing the sexual forms of the parasite, called gametocytes. Once inside the mosquito's stomach, the male cell extends and then detaches flagellum-like structures which migrate to and then fertilize the female cells. The fertilized cells penetrate the wall of the mosquito's stomach and there grow into oocysts containing filament-like structures called sporozoites. Upon reaching maturity, the oocysts rupture, releasing up to several hundred thousand sporozoites to migrate throughout the mosquito's body cavity. Some of these sporozoites reach the salivary glands and there remain dormant until injected into man. For *P. falciparum* this anopheline phase of the life cycle lasts between 7 and 14 days.

The sporozoites disappear from the peripheral blood within 30 minutes after injection into man, initiating the exoerythrocytic stage of the life cycle. The parasites develop in the liver parenchymal cells and reach maturity after about six days. The mature parasites are about 60 μm in their longest diameter and release approximately 40,000 merozoites into the peripheral blood to invade erythrocytes. At this stage of development parasites can first be detected in the blood. A minimum of 10 parasites per mm^3 is normally required for detection by ordinary microscopic examination of a thick blood smear.

The invasion of red blood cells by merozoites initiates the production of the next stage of asexual reproduction, trophozoites. In about 48 hours, each trophozoite releases from 8 to 24 new merozoites for further invasion of red blood cells. After several generations of this process, the parasitized blood cells release male and female gametocytes capable of infecting anopheles mosquitos. A series of trophozoite-gametocyte waves typically follows. As the gametocyte count rises, the trophozoite count falls, and clinical improvement or remission of symptoms frequently occurs. Parasite

counts in falciparum malaria fluctuate markedly. They often show alternating high and low densities on successive days.

2.2 *FIELD SURVEYS*

For at least sixty years there have been international efforts to control malaria in Africa, Asia, and the Americas. Important ingredients in the planning of antimalarial programs are baseline data and quantitative estimates of the key factors governing transmission. A good quantitative understanding of the natural course of malaria requires longitudinal field surveys of the human population synchronized with mosquito surveys. Such investigations are very costly and difficult to organize and implement. Despite the widespread recognition that such studies are essential to estimate human incidence rates of infection and recovery and entomological inoculation rates, malaria surveys on humans from 1920 to 1970 were almost exclusively cross-sectional prevalence surveys. For a review of malaria prevalence surveys see [26].

The WHO surveys in Garki are the first to make possible direct estimates of the age-specific rates, r_A and ρ_A, where A may be either apparent new infection or apparent recovery. In the Garki project, sixteen villages were surveyed every ten weeks from the end of the wet season in November 1970 to the end of the dry season in May 1972. There were no attempts at treatment or control of malaria during this period. The surveys are therefore called baseline surveys. The surveys aimed at complete coverage of each village and included the collection of a thick blood film, linked by a code number to the person's identity. Blood films were stained with Giemsa stain and examined, under oil immersion, with 7× oculars and 100× objective, for 200 fields. Each person was classified as positive for *P. falciparum* parasitemia if trophozoites and/or gametocytes were observed in any of the 200 fields.

Data on the mosquito vector were collected in some of the 16 villages every 5 weeks in the dry season and every 2 weeks in the wet season. Man-biting rates and entomological inoculation rates, defined as the average number of sporozoite-positive bites per man per night, were obtained using human bait. For further details on the entomological surveys in Garki, see [21].

Entomological survey data provide estimates of inoculation rates, which have appeared as parameters in proposed models of malaria transmission dating from 1911 [24] to the present. Despite the substantial theoretical literature on the epidemiology of transmission [18,20,19,2], the lack of any data comparable to those of the Garki surveys has prevented systematic field testing of these proposals. Dietz et al. [10] attempted to model the

transmission of malaria using data from the baseline surveys in Garki. They explicitly took account of the effect of immunity on transmission, of superinfection, and of recovery rates (expected number of recoveries per unit time at time s per person positive at time s).

The development of an integrated stochastic model of the infection histories in a human population, coupled to a model of the dynamics of the mosquito vector, lies in the future. Our aim here is much more modest. Using a time series of two-wave panel surveys (Sec. 4) we ascertain whether or not the observations could have been generated by continuous-time Markov chains or certain mixtures of such chains. Having accepted a restricted class of processes (e.g. the time-homogeneous Markov chains) as consistent with the panel data, we estimate the rates r_A and ρ_A within this class. We then interpret their age- and season-specific values in the context of malaria dynamics in an unprotected population.

Although we analyze only the baseline surveys in Garki, the age- and season-specific rates ρ_A and r_A can also be compared with the corresponding quantities estimated in some of the same villages during $1\frac{1}{2}$ years of interventions. Though we shall not attempt to use incidence rates to assess the impact of interventions, Sec. 3–6 sketch strategies basic to such evaluations. For data from the intervention and follow-up phases in Garki, see [21].

3. SAMPLING ASPECTS OF THE SIMPLEST MARKOV EMBEDDING PROBLEM

3.1. *EMBEDDABILITY CRITERIA*

Consider a two-parameter family of finite $r \times r$ stochastic matrices $P(s,t)$, $0 \leqslant s \leqslant t < +\infty$, in which (a) each element $p_{ij}(s,t)$ of $P(s,t)$ is continuous in (s,t), (b) $P(s,t) = I$ if and only if $s = t$, and (c) $P(s,t) = P(s,u)P(u,t)$ whenever $s \leqslant u \leqslant t$.

Such a family $P(s,t)$ may be thought of as describing a nonstationary continuous-time finite Markov chain. $p_{ij}(s,t)$ gives the transition probability from state i at time s to state j at time t. We call a finite stochastic matrix P embeddable in a continuous-time Markov chain if there exists a two-parameter family of stochastic matrices $P(s,t)$ satisfying (a), (b), (c), and (d) $P(0,1) = P$.

Goodman [12] proved that any embeddable matrix P could, by a change of time scale, be embedded in a continuous two-parameter family of transition matrices $P(s,t)$ satisfying, for almost all s and t,

$$\frac{\partial P(s,t)}{\partial t} = P(s,t)Q(t), \qquad P(t,t) = I \tag{3.1a}$$

and

$$\frac{\partial P(s,t)}{\partial s} = -Q(s)P(s,t), \qquad P(s,s)=I, \tag{3.1b}$$

where $Q(s)$ is a bounded measurable function such that for all $s \geqslant 0$,

$$Q(s) \in \mathbf{Q} = \left\{ Q \,|\, q_{ii} \leqslant 0,\ q_{ij} \geqslant 0,\ i \neq j,\ \sum_{j=1}^{r} q_{ij} = 0 \right\}.$$

$\mathbf{Q}$ is the set of all "intensity matrices." Equations (3.1a) and (3.1b) are known as the Kolmogorov forward and backward equations, respectively.

THEOREM [12]

A 2×2 *stochastic matrix* P *is embeddable in a continuous-time Markov chain if and only if* $\operatorname{trace} P > 1$.[1]

This simple and very useful characterization was established by D. G. Kendall [16, p. 15] for homogeneous chains, i.e., those for which $P(s,t) \equiv P(t-s)$, and extended to inhomogeneous chains by Goodman [12].

Although a 2×2 stochastic matrix P with $\operatorname{trace} P > 1$ can be embedded in uncountably many inhomogeneous Markov chains, such a matrix is embeddable in a unique homogeneous chain. To see this, observe that the transition matrices for homogeneous chains solve (3.1a) and (3.1b) with $Q(s) = \text{constant}$. The unique solution is

$$P(s,t) = P(t-s) = e^{(t-s)Q}, \qquad 0 \leqslant s \leqslant t < +\infty. \tag{3.2}$$

Thus

$$P = e^{Q} \qquad \text{for some} \quad Q \in \mathbf{Q}, \tag{3.3}$$

or

$$Q = \log P = \frac{\log(p_{11}+p_{22}-1)}{p_{11}+p_{22}-2} \begin{pmatrix} p_{11}-1 & 1-p_{11} \\ 1-p_{22} & p_{22}-1 \end{pmatrix}, \tag{3.4}$$

where

$$P = \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix}$$

[1] Goodman proved that $\det P > 0$ is a necessary condition for embeddability of a stochastic matrix P of any order. For 2×2 matrices, this is equivalent to $\operatorname{trace} P > 1$. Sufficiency is an immediate consequence of (3.4), since $\operatorname{trace} P > 1$ implies that P is embeddable in a time-homogeneous chain.

with $p_{11}+p_{22}=\text{trace}\,P>1$. Since the principal branch of the logarithm in (3.4) is the only determination of $\log P$ which is a member of $\mathbf{Q}$, this establishes the uniqueness of embeddability for 2×2 homogeneous chains. The uniqueness breaks down and the embeddability criteria become more complicated for matrices of homogeneous chains of order 3 or more [9,28].

In practical applications where it is natural to consider embedding a stochastic matrix P in a continuous-time family of such matrices, P must be estimated from sampled data. The estimated matrix $\hat{P}$ may fail to be embeddable even when P is. Moreover, $\log \hat{P}$ is then only an estimate of $\log P$, and it is important to know how much variation should be expected in $\log \hat{P}$ given P and the sample size. We now develop a formal test for embeddability, assuming binomial sampling, based on the criterion trace $P>1$.

3.2. *TWO-WAVE PANEL DATA AND BINOMIAL SAMPLING*

For individuals evolving independently and observed at two points in time, we may describe their location in one of two states—e.g., not infected or infected in malaria parasitemia surveys—according to the following probability model.

Let n_{i+} be the number of individuals in state i at the initial observation time $t=0$, $i=1,2$, and introduce the independent, identically distributed random variables $Y_1,\ldots,Y_{n_{1+}}$ where

$$Y_i=\begin{cases}1 & \text{with probability } p_{11},\\ 0 & \text{otherwise.}\end{cases}$$

Here p_{11} is the *a priori* unknown conditional probability $\text{Prob}(w: w(\Delta)=1 | w(0)=1)$, where Δ is the time when the second wave of panel data is collected and $\{w(t),\ t\geqslant 0\}$ is a parasitemia history, identified as a sample path of some stochastic process on the state space $\{1,2\}$.

Let $Z_1,\ldots,Z_{n_{2+}}$ be independent, identically distributed random variables, which are also independent of $Y_1,\ldots,Y_{n_{1+}}$, and set

$$Z_i=\begin{cases}1 & \text{with probability } p_{22},\\ 0 & \text{otherwise.}\end{cases}$$

Here $p_{22}=\text{Prob}(w: w(\Delta)=2 | w(0)=2)$.

If

$$n_{11}=\sum_{i=1}^{n_{1+}} Y_i$$

and

$$n_{22} = \sum_{j=1}^{n_{2+}} Z_j,$$

n_{ii} may be interpreted as the number of individuals observed in state i at time $t = \Delta$ who were also observed in state i at time $t = 0$.

The conditional probabilities p_{ii}, $i = 1, 2$, are not assumed to arise from any restricted class of stochastic processes such as the continuous-time Markov chains. The above formulation assumes only that individuals evolve independently and defers the specification of the underlying dynamics. In our procedure, a specification of the dynamics follows tests of hypotheses applied to the data. A similar approach is taken by Anderson and Goodman [1], who test multinomial frequencies—estimated from multiwave panel data—for consistency with conditional probabilities arising in Markov processes of order $1, 2, \ldots,$. However, the sampling theory of the embeddability criterion for continuous-time processes has not been systematically discussed before.

By contrast to this approach, it is usually assumed at the outset that the observations are generated by independent realizations of a continuous-time Markov chain, or mixture of Markov chains, semi-Markov processes, etc. (e.g. [5], [15], and their references).

In order to create a formal test of hypotheses based on the criterion $\operatorname{trace} P > 1$, we introduce the null hypothesis

$$H_0 : \operatorname{trace} P > 1,$$

and the alternative hypothesis

$$H_1 : \operatorname{trace} P \leqslant 1.$$

We propose a decision rule based on δ_1 and δ_2 to be defined as follows:

If $\operatorname{trace} \hat{P} > 1 + \delta_1$, then accept H_0.

If $1 - \delta_2 \leqslant \operatorname{trace} \hat{P} \leqslant 1 + \delta_1$, then the evidence is inadequate to distinguish between H_0 and H_1.[2]

If $\operatorname{trace} \hat{P} < 1 - \delta_2$, then accept H_1.

---

[2]This is an instance of a three-decision rule of the kind discussed by Lehman [17, p. 549]. The more common two-decision rules, which require acceptance or rejection of a proposed hypothesis, are excessively stringent. They fail to formalize the Scotch verdict "not proven" when values of a test statistic lie close to the boundary of the set of values that are consistent with the hypothesized models.

Given error probabilities specified *a priori* as $\alpha_1 = \text{Prob}_{H_0}(\text{reject } H_0)$ and $\alpha_2 = \text{Prob}_{H_1}(\text{accept } H_0)$, we define δ_2 by

$$\alpha_1 = \sup_{\mathbf{p}: p_{11}+p_{22}>1} \text{Prob}_{\mathbf{p}}(\text{trace}\,\hat{P} < 1-\delta_2) \tag{3.5}$$

and define δ_1 by

$$\alpha_2 = \sup_{\mathbf{p}: p_{11}+p_{22}<1} \text{Prob}_{\mathbf{p}}(\text{trace}\,\hat{P} > 1+\delta_1). \tag{3.6}$$

Here $\boldsymbol{p} = (p_{11}, p_{22})$,

$$\text{trace}\,\hat{P} = \frac{1}{n_{1+}} \sum_{i=1}^{n_{1+}} Y_i + \frac{1}{n_{2+}} \sum_{j=1}^{n_{2+}} Z_j,$$

and the probabilities in (3.5) and (3.6) are computed from the independent binomial distributions which describe $\sum_{i=1}^{n_{1+}} Y_i$ and $\sum_{j=1}^{n_{2+}} Z_j$. For example, using $\lfloor x \rfloor$ to denote the integer part of a nonnegative real x,

$$\begin{aligned}
&\text{Prob}_{\mathbf{p}}(\text{trace}\,\hat{P} < 1-\delta_2) \\
&= \sum_{k=0}^{\lfloor n_{1+}n_{2+}(1-\delta_2) \rfloor} \sum_{(l,m): l+m=k} \binom{n_{1+}}{\lfloor l/n_{2+} \rfloor} p_{11}^{\lfloor l/n_{2+} \rfloor} (1-p_{11})^{n_{1+}-\lfloor l/n_{2+} \rfloor} \\
&\quad \times \binom{n_{2+}}{\lfloor m/n_{1+} \rfloor} p_{22}^{\lfloor m/n_{1+} \rfloor} (1-p_{22})^{n_{2+}-\lfloor m/n_{1+} \rfloor}.
\end{aligned} \tag{3.7}$$

Although (3.7) and the analogous formula for $\text{Prob}_{\mathbf{p}}(\text{trace}\,\hat{P} > 1+\delta_1)$ are unwieldy for solving (3.5) and (3.6) for δ_i, the central limit theorem yields

$$\alpha_1 \approx \Phi\left(\frac{-2\delta_2}{(n_{1+}^{-1} + n_{2+}^{-1})^{1/2}} \right), \tag{3.8}$$

$$\alpha_2 \approx 1 - \Phi\left(\frac{2\delta_1}{(n_{1+}^{-1} + n_{2+}^{-1})^{1/2}} \right). \tag{3.9}$$

As Table 1 indicates, these approximations are quite accurate for samples where $\min(n_{1+}, n_{2+}) = 10$.

TABLE 1

Comparison of Asymptotic and Small-Sample Evaluations of δ_2[a]

Sample sizes		Exact value of δ_2				
n_{1+}	n_{2+}	.050	.100	.150	.200	.250
10	11	Not computed	.087 (.345)	.164 (.226)	.185 (.198)	.273 (.106)
50	51	.057 (.284)	.100 (.158)	.163 (.051)	.213 (.016)	.308 (.001)
51	74	.049 (.294)	.102 (.131)	.145 (.055)	.189 (.019)	.241 (.004)
39	62	.049 (.317)	.100 (.164)	.151 (.070)	.188 (.033)	.246 (.008)
44	407	.053 (.254)	.100 (.104)	.147 (.032)	.195 (.007)	.245 (.001)

[a]The exact value of δ_2 given at the top of each column is to be compared with the asymptotic value of δ_2 given as the upper value in each cell. The asymptotic value of δ_2 solved (3.8), given the sample sizes n_{1+} and n_{2+} for the row and the value of α_1 in parentheses in the same cell. α_1 is computed by estimating the supremum in (3.5) over the set of (p_{11},p_{22}) pairs

$\{(.5,.5), (.9,.1), (.9,.2), (.9,.4), (.9,.5), (.9,.6), (.9,.8), (.9,.9), (.8,.2), (.8,.4), (.8,.5), (.8,.6), (.8,.8), (.6,.4), (.6,.5)\}$.

For $n_{1+}=10, n_{2+}=11$, the probabilities are computed exactly from the binomial distribution. For the remaining sample sizes, the probabilities are the sample proportions in 1000 Monte Carlo evaluations of $B(n_{1+},p_{11})/n_{1+} + B(n_{2+},p_{22})/n_{2+}$ which are $< 1-\delta_2$, where $B(n_{i+},p_{ii})$, $i=1,2$, are independent binomial random variates with parameters n_{i+},p_{ii}.

This embeddability test is intended to assess whether 2-wave, 2-state panel data are consistent with some continuous-time Markov-chain model. However, interpreting an acceptance or rejection of H_0 is not simple. For many non-Markovian processes with transition matrices $P=P(0,\Delta)$, it is true that trace $P>1$ for many or all matrices P. Such processes are indistinguishable from the continuous-time Markov chains on the basis of two-wave panel data. Thus when H_0 is accepted in the formal test described above and continuous-time Markov models are utilized for making further inferences (e.g. determining conversion and recovery rates from *P. falciparum* parasitemia), such models are tentative and require further validation utilizing 3 or more waves of panel data. Estimates of event rates may be biased if alternative but empirically indistinguishable processes

actually generate the data. For example, transition matrices for mixtures of continuous-time Markov chains

$$P_\mu(0,t) = \int_{\mathbf{Q}} P^{(Q)}(0,t)\, d\mu(Q), \tag{3.10}$$

where μ is an arbitrary probability measure on the space $\boldsymbol{Q}$ of measurable intensity matrices Q, must satisfy $\operatorname{trace} P_\mu(0,t) > 1$. Here $P^{(Q)}(0,t)$ are transition probabilities for a 2-state continuous-time Markov chain with intensity matrix function $Q(t)$, $t \geqslant 0$. In Sec. 7.2 we discuss biases in event-rate estimates which can arise from the use of a simple Markov-chain model rather than a nondegenerate mixture (3.10).

In contrast with the weakness of inferences associated with the acceptance of H_0, a rejection of H_0 leads to strong conclusions. If $\operatorname{trace} \hat{P}(0,\Delta) < 1 - \delta_2$, then no mixture of the form (3.10)—degenerate or not—could describe the data. Surprisingly, this procedure eliminates large classes of models from contention as candidates to describe an evolving process, without directly testing dynamic characteristics such as the Markov property. Statistical tests based on characterizations of the conditional probabilities that can be generated by special models exhibit their greatest power as tools for model rejection.

Our decision rule allows the possibility of no decision when $\operatorname{trace} \hat{P}$ is sufficiently close to 1. The intuitively appealing properties of the probability of no decision, $\operatorname{Prob}_{\mathbf{p}}(1 - \delta_2 \leqslant \operatorname{trace} \hat{P} \leqslant 1 + \delta_1)$, for the embeddability test are: (i) For each (p_{11}, p_{22}) such that $p_{11} + p_{22} \neq 1$, and each (n_{1+}, n_{2+}), if $\alpha_i, i = 1,2$ decrease, then the probability of no decision increases. (ii) For each $\alpha_i \in (0, \frac{1}{2})$ and p_{ii} such that $p_{11} + p_{22} \neq 1$, if $\min(n_{1+}, n_{2+})$ increases, then the probability of no decision decreases. (iii) For each $\alpha_i \in (0, \frac{1}{2})$ and (n_{1+}, n_{2+}), if $|p_{11} + p_{22} - 1|$ increases, then the probability of no decision decreases.

To provide further insight, Table 2 illustrates numerically the relationships among α_i, $\mathbf{p}$, (n_{1+}, n_{2+}), and the probability of no decision.

3.3. *HOMOGENEOUS CHAINS*

a. Estimation of Intensity Matrices. Following the acceptance of H_0 in the embeddability test of Sec. 3.2, the simplest stochastic models which describe two-wave panel data are the homogeneous Markov chains. We present the well-known maximum-likelihood estimator $\hat{Q}$ of the unique 2×2 intensity matrix

$$Q = \begin{pmatrix} -q_1 & q_1 \\ q_2 & -q_2 \end{pmatrix}$$

TABLE 2

Probabilities of No Decision and Critical Values of $\delta = \delta_1 = \delta_2$ [a]

Sample size		Main diagonal (p_{11}, p_{22}) of P				
n_{1+}	n_{2+}	(.9,.2)	(.9,.9)	(.8,.4)	(.6,.6)	δ
		Probability of no decision using binomial distribution				
10	11	.995	.015	.936	.904	.5082
		.937	.003	.820	.796	.3594
		.912	.001	.693	.638	.2801
		Probability of no decision using normal approximation				
10	11	.996	.013	.942	.925	.5082
		.951	0	.789	.767	.3594
		.869	0	.651	.633	.2801
25	51	.979	0	.771	.759	.2839
		.867	0	.503	.502	.2008
		.732	0	.349	.357	.1565
100	501	.742	0	.082	.088	.1274
		.407	0	.018	.020	.0901
		.240	0	.006	.008	.0702
500	501	.118	0	0	0	.0735
		.016	0	0	0	.0520
		.004	0	0	0	.0405

[a]Values less than 0.0005 are shown as 0. In each cell, the upper, middle, and lower entries correspond to $\alpha_1 = \alpha_2 = \alpha = 0.01, 0.05, 0.10$. For each α and sample size, δ was obtained from (3.8). The probability of no decision, $\text{Prob}(1 - \delta < \text{trace}\,\hat{P} < 1 + \delta)$, was obtained from binomial probabilities for $n_{1+} = 10$, $n_{2+} = 11$ and from a normal approximation for all sample sizes.

which specifies the chain's transition probabilities by (3.2). We also assess the variability of $\hat{Q}$.

Assuming binomial sampling,

$$\hat{P}(0, \Delta) = (n_{ij}/n_{i+}), \qquad i,j = 1,2, \tag{3.11}$$

is the maximum-likelihood estimator of $P(0, \Delta)$. If (3.11) satisfies $\text{trace}\,\hat{P} > 1$, then $\Delta^{-1} \log \hat{P}$ is the maximum-likelihood estimator of Q in the representation (3.2) with $\Delta = t - s$.

Thus the maximum-likelihood estimators of q_1 and q_2 for embeddable matrices $\hat{P}$ are

$$\Delta \hat{q}_i = \frac{(1 - \hat{p}_{ii}) \log(\text{trace}\,\hat{P} - 1)}{\hat{p}_{11} + \hat{p}_{22} - 2}, \qquad i = 1,2. \tag{3.12}$$

These are consistent but not unbiased, asymptotically normal estimators of Q.

b. Variability of Estimators. The variance-covariance matrix $(\mathrm{cov}(\hat{q}_i, \hat{q}_j))$ is given, asymptotically, by the inverse of the Fisher information matrix:

$$I^{-1} = -\begin{bmatrix} \dfrac{\partial^2 \log L}{\partial q_1^2} & \dfrac{\partial^2 \log L}{\partial q_1 \partial q_2} \\ \dfrac{\partial^2 \log L}{\partial q_1 \partial q_2} & \dfrac{\partial^2 \log L}{\partial q_2^2} \end{bmatrix}^{-1}, \tag{3.13}$$

where

$$\log L = \sum_{i=1}^{2} \sum_{j=1}^{2} n_{ij} \log P_{ij} + \sum_{i=1}^{2} \log \binom{n_{i+}}{n_{ii}}$$

and

$$P_{ij}(0, \Delta) = \frac{(1 - e^{-(q_1+q_2)\Delta}) q_i}{q_1 + q_2},$$
$$P_{ii}(0, \Delta) = 1 - P_{ij}(0, \Delta), \qquad i \neq j.$$

For embeddable matrices $\hat{P}$, the explicit formulas are

$$\begin{aligned} \mathrm{Var}(\hat{q}_1) &= (S_1[\hat{p}_{12}V + \hat{p}_{21}U][\hat{p}_{12}V + \hat{p}_{21}U]^* \\ &\quad + S_2[-\hat{p}_{12}V + \hat{p}_{12}U][-\hat{p}_{12}V + \hat{p}_{12}U]^*)(\hat{p}_{12} + \hat{p}_{21})^{-4}, \\ \mathrm{Var}(\hat{q}_2) &= (S_1[-\hat{p}_{21}V + \hat{p}_{21}U][-\hat{p}_{21}V + \hat{p}_{21}U]^* \\ &\quad + S_2[\hat{p}_{21}V + \hat{p}_{12}U][\hat{p}_{21}V + \hat{p}_{12}U]^*)(\hat{p}_{12} + \hat{p}_{21})^{-4}, \\ \mathrm{Cov}(\hat{q}_1, \hat{q}_2) &= -(S_1[\hat{p}_{12}V + \hat{p}_{21}U][-\hat{p}_{21}V + \hat{p}_{21}U]^* \\ &\quad + S_2[-\hat{p}_{12}V + \hat{p}_{12}U][\hat{p}_{21}V + \hat{p}_{12}U]^*)(\hat{p}_{12} + \hat{p}_{21})^{-4}, \end{aligned} \tag{3.14}$$

where * is the complex conjugate and

$$U = -\log(\mathrm{trace}\,\hat{P} - 1),$$
$$V = \frac{\hat{p}_{12} + \hat{p}_{21}}{\mathrm{trace}\,\hat{P} - 1},$$
$$S_i = \frac{\hat{p}_{ij}(1 - \hat{p}_{ij})}{n_{i+}}, \qquad j \neq i.$$

To develop guidelines about the size of $\min(n_{1+}, n_{2+})$ which ensures that I^{-1} is a good approximation to $(\mathrm{cov}(\hat{q}_i, \hat{q}_j))$, we carried out some exploratory Monte Carlo calculations. Without describing the details here, we offer several generalizations.

(i) For small samples—$\max(n_{1+}, n_{2+}) \leqslant 20$—we have $\mathrm{Var}(\hat{q}_i) \approx 2 \times$ [asymptotic approximations of $\mathrm{Var}(\hat{q}_i)$].

(ii) If $\min(n_{1+}, n_{2+}) \leqslant 100$, compute $\mathrm{cov}(\hat{q}_i, \hat{q}_j)$ from the binomial distribution or a Monte Carlo approximation based on binomial sampling, as described below.

(iii) If $\min(n_{1+}, n_{2+}) > 1000$, then the asymptotic maximum-likelihood approximation is very good.

(iv) If $100 < \min(n_{1+}, n_{2+}) \leqslant 1000$, we recommend Monte Carlo evaluation of $\mathrm{cov}(\hat{q}_i, \hat{q}_j)$. This suggestion is guided by the comparison (Table 5) of Monte Carlo calculations and maximum-likelihood approximations to $\mathrm{cov}(\hat{q}_i, \hat{q}_j)$ in the WHO malaria data, where many of the tables have sample sizes in the range $100 < n_{i+} < 1000$.

For the Monte Carlo calculations of $\mathrm{cov}(\hat{q}_i, \hat{q}_j)$ from two-wave panel data, we generated 1000 stochastic matrices from binomial samples of sizes n_{1+} and n_{2+} with parameters $\hat{p}_{ii} = n_{ii}/n_{i+}$, $i = 1, 2$, respectively. The binomial variates were generated from a subroutine of the International Mathematical and Statistical Library, version 5, resident at the Computing Center of the City University of New York. We then computed 1000 values of $\hat{q}_i$, $i = 1, 2$, using (3.12). Let $(\hat{q}_i)_k$ denote the value of (3.12) computed from the kth matrix, $k = 1, \ldots, 1000$. Our Monte Carlo estimates are defined by

$$E_{\mathrm{MC}}(\hat{q}_i) = 0.001 \sum_{k=1}^{1000} (\hat{q}_i)_k,$$

$$\mathrm{Var}_{\mathrm{MC}}(\hat{q}_i) = 0.001 \sum_{k=1}^{1000} (\hat{q}_i)_k^2 - [E_{\mathrm{MC}}(\hat{q}_i)]^2,$$

$$\mathrm{Cov}_{\mathrm{MC}}(\hat{q}_1, \hat{q}_2) = 0.001 \sum_{k=1}^{1000} (\hat{q}_1)_k (\hat{q}_2)_k - E_{\mathrm{MC}}(\hat{q}_1) E_{\mathrm{MC}}(\hat{q}_2).$$

The Monte Carlo calculations were used instead of the exact sampling distribution of $\hat{p}_{ii}$, $i = 1, 2$, since computation based on the exact distribution is prohibitively long except for very small sample sizes [$\max(n_{1+}, n_{2+}) \leqslant 15$].

4. A TIME SERIES OF TWO-WAVE PANELS

Table 3 (adapted from [4, Table 1] lists the number n_{ij} of persons in state i at one survey and in state j at the next, for five pairs of successive surveys 3 to 4, ..., 7 to 8. We define a positive person as being in state 2. A negative person, one without detectable parasitemia, is in state 1. We give ages in years of age at survey 1. These surveys span one year. The variation in numbers of persons in the consecutive 2×2 tables for a single age class

TABLE 3

Malaria Parasitology Transition Data: Observed Numbers n_{ij} of Persons in State i at First Wave of Two-Wave Panel Survey and State j at the Second Wave[a]

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

[a]For each two-wave panel survey and age class, the entries in the 2×2 table are $\begin{bmatrix} n_{11} & n_{12} \\ n_{21} & n_{22} \end{bmatrix}$, where 1 = negative, 2 = positive. Thus of individuals 5 to 8, 52 people uninfected at survey 5 were infected at survey 6, and 68 infected at survey 5 were uninfected at survey 6. Source: Bekessy et al. [4].

results primarily from some people being absent from a village on the survey dates. Among infants, new births and aging from <1 to 1–4 contribute to the variation.

We apply the Markov embeddability test of Sec. 3.2 to each 2×2 array in Table 3. As indicated in Table 4, all but three of the 2×2 arrays are consistent with some continuous-time Markov process. The three exceptions justify no decision. One possible explanation for these exceptions is measurement error due to the daily fluctuations in parasite density in an infected individual (see Sec. 7.1).

Table 5 gives the maximum-likelihood estimates $\hat{q}$, $i=1,2$, of the parameters in the unique time-homogeneous chain that embeds each Markov-embeddable 2×2 table together with the variance-covariance matrix computed via the Monte Carlo method (Sec. 3.3) and the inverse of the Fisher information matrix (3.13).

We interpret $1/\hat{q}_1$ as the expected duration of a spell without patent parasitemia that is initiated during the 10-week period between the successive surveys. Similarly, we interpret $1/\hat{q}_2$ as the expected duration of a spell

TABLE 4

Embeddability Tests for Malaria Infection and Recovery Data in Table 3[a]

Surveys	Age class	Sample sizes n_{1+}	Sample sizes n_{2+}	δ for $\alpha=0.01$	δ for $\alpha=0.05$	δ for $\alpha=0.10$	trace $\hat{P}$
3–4	<1	76	54	.207	.146	.114	1.580
	1–4	59	494	.160	.113	.088	1.551
	5–8	74	542	.144	.102	.079	1.258
	9–18	181	357	.106	.075	.058	1.370
	19–28	478	191	.100	.070	.055	1.341
	29–43	1011	270	.080	.056	.044	1.197
	44+	622	160	.103	.073	.057	1.200
4–5	<1	87	30	.246	.174	.136	1.041[b]
	1–4	91	490	.133	.094	.073	1.157
	5–8	96	611	.128	.090	.070	1.138
	9–18	237	440	.094	.066	.052	1.283
	19–28	541	250	.089	.063	.049	1.238
	29–43	1099	353	.071	.050	.039	1.149
	44+	640	184	.097	.069	.054	1.094[b]
5–6	<1	39	62	.238	.168	.131	1.108[b]
	1–4	50	484	.173	.122	.095	1.212
	5–8	70	649	.146	.103	.081	1.152
	9–18	169	537	.103	.073	.057	1.240
	19–28	433	356	.083	.059	.046	1.189
	29–43	912	569	.062	.044	.034	1.127
	44+	494	311	.084	.060	.046	1.106
6–7	<1	39	79	.228	.161	.125	1.566
	1–4	45	454	.182	.129	.100	1.285
	5–8	71	550	.147	.104	.081	1.190
	9–18	147	380	.113	.080	.062	1.306
	19–28	345	300	.092	.065	.051	1.232
	29–43	842	462	.067	.048	.037	1.171
	44+	504	243	.091	.064	.050	1.147
7–8	<1	69	71	.197	.139	.108	1.674
	1–4	44	407	.185	.131	.102	1.426
	5–8	80	534	.139	.099	.077	1.239
	9–18	179	322	.108	.077	.060	1.419
	19–28	436	179	.103	.073	.057	1.328
	29–43	1014	262	.081	.057	.044	1.219
	44+	636	136	.110	.078	.061	1.161

[a]The decision rule in Sec. 3.2 was implemented here by setting $\alpha_1 = \alpha_2$ and computing δ from (3.8).

[b]Here trace $\hat{P} < 1 + \delta$ for the δ corresponding to at least one value of α. Therefore, with this value of α, the decision rule would recommend no decision regarding embeddability. There were no cases where trace $\hat{P} < 1$.

TABLE 5

Malaria Conversion Intensity $\hat{q}_1$ and Recovery Intensity $\hat{q}_2$ and Their Variances and Covariance Estimated by Maximum-Likelihood and Monte Carlo Simulation

		Intensity[b] (day^{-1})		Maximum-likelihood estimate[c] (10^{-8} day^{-2})			Monte Carlo estimate[d] (10^{-8} day^{-2})		
Surveys[a]	Age groups	$\hat{q}_1$	$\hat{q}_2$	$\mathrm{Var}(\hat{q}_1)$	$\mathrm{Cov}(\hat{q}_1,\hat{q}_2)$	$\mathrm{Var}(\hat{q}_2)$	$\mathrm{Var}(\hat{q}_1)$	$\mathrm{Cov}(\hat{q}_1,\hat{q}_2)$	$\mathrm{Var}(\hat{q}_2)$
3–4	<1	.0038	.0042	103.8	36.8	162.2	118.9	41.1	170.8
	1–4	.0070	.0018	246.5	1.7	9.0	250.9	15.1	9.0
	5–8	.0171	.0029	880.2	86.5	23.8	1023.0	108.1	28.1
	9–18	.0090	.0057	151.4	43.3	49.7	157.9	43.3	49.7
	19–28	.0050	.0108	36.8	43.3	183.9	36.8	43.3	177.3
	29–43	.0059	.0179	36.8	84.3	374.1	43.3	99.5	413.1
	44+	.0054	.0183	51.9	12.5	625.0	56.2	138.4	670.4
4–5	<1	.0323	.0085	37,439.0	14,072.0	5813.0	29,101.0	$9627.0-2183i$	4033.0[e]
	1–4	.0220	.0018	1,173.0	78.9	14.8	1,345.0	80.5	14.8
	5–8	.0233	.0021	1,359.0	100.2	18.1	1,616.0	121.6	19.7
	9–18	.0129	.0033	180.7	34.5	23.1	179.2	37.8	26.3
	19–28	.0089	.0095	83.8	78.9	157.8	87.1	85.5	166.0
	29–43	.0102	.0141	98.6	164.4	292.6	100.3	142.9	295.9
	44+	.0125	.0178	455.2	708.4	1330.0	950.0	$1496.0-14.8i$	2604.0[e]
5–6	<1	.0245	.0030	5377.0	812.4	257.6	18,262.0	$2887.0-216.4i$	632.5[e]
	1–4	.0175	.0017	1248.0	76.2	13.7	1,463.0	102.1	16.8
	5–8	.0204	.0029	1542.0	1574.0	28.9	2,164.0	240.8	42.7[e]
	9–18	.0123	.0053	285.0	73.2	42.7	321.6	80.8	45.7
	19–28	.0094	.0112	124.9	109.7	178.3	150.9	143.3	216.4
	29–43	.0089	.0166	88.4	129.6	278.9	106.7	160.0	342.9[e]
	44+	.0088	.0190	195.1	326.2	756.0	307.9	$539.6-1.5i$	1186.0[e]
6–7	<1	.0049	.0026	235.4	34.6	64.1	232.0	46.7	72.7
	1–4	.0149	.0016	986.8	58.9	12.1	1143.0	67.5	12.1
	5–8	.0186	.0032	1222.0	141.9	32.9	1670.0	193.9	39.8
	9–18	.0093	.0063	207.8	67.5	60.6	211.2	65.8	62.3
	19–28	.0047	.0145	51.9	65.8	219.9	57.1	74.4	245.8
	29–43	.0042	.0191	24.2	50.2	244.1	24.2	55.4	277.0
	44+	.0042	.0211	45.0	114.3	593.8	51.9	138.5	692.5
7–8	<1	.0015	.0041	38.8	12.2	106.1	38.8	12.2	114.3
	1–4	.0111	.0011	597.9	22.4	6.1	642.9	26.5	8.2
	5–8	.0181	.0023	875.5	71.4	18.4	1016.0	87.8	20.4
	9–18	.0071	.0054	104.1	30.6	46.9	112.2	32.7	42.9
	19–28	.0044	.0115	34.7	42.9	208.2	34.7	44.9	212.2
	29–43	.0046	.0171	22.4	55.1	308.2	24.5	61.2	318.4
	44+	.0053	.0208	63.3	189.7	1012.0	73.5	232.6	1208.0

[a]There were 68 days between surveys 3 and 4, 78 between 4 and 5, 81 between 5 and 6, 76 between 6 and 7, and 70 between 7 and 8.

[b]Taken from [4, Table 1].

[c]Calculated from (3.14).

[d]Based on 1000 trials for each age group in each survey.

[e]Note substantial difference between the Monte Carlo and asymptotic maximum-likelihood estimates of at least one variance or covariance term.

of patent parasitemia that is initiated during the 10-week period between successive surveys. It is usual in the Garki district that persons have multiple spells of patent parasitemia and freedom from parasitemia. Initiation of a spell of freedom from patent parasitemia may be interpreted as an apparent recovery without treatment from *P. falciparum* parasitemia. We use the word "apparent" because the transition from "positive" to "negative" may result either from a true recovery (getting rid of the infection) or from going into a state of latency (with very few or no parasites in the peripheral blood). The Garki data do not allow one to separate these two kinds of transitions from positive to negative.

All of the two-wave panels in Table 3 are embeddable. In the Garki project there are some nonembeddable two-wave panels in the intervention and follow-up phases. For example, infants in village clusters 5 and 7 at surveys 21 and 22 (wet season, 1975) have transition counts and probabilities given, respectively, by

$$(n_{ij}) = \begin{pmatrix} 68 & 28 \\ 17 & 4 \end{pmatrix} \quad \text{and} \quad \hat{P} = \begin{pmatrix} .7083 & .2917 \\ .8095 & .1905 \end{pmatrix}. \tag{4.1}$$

Thus

$$\operatorname{trace} \hat{P} = 0.8988 < 1.$$

We have chosen not to discuss the tables in the intervention and follow-up phases—such as (4.1)—that reject the class of continuous-time Markov models, since we are then confronted with the (yet unresolved) issue of providing simple non-Markovian models that describe those data.

The extent of nonembeddable two-wave, two-state panel data generally is impossible to ascertain because of the very limited experience with stochastic modeling of such data. The effort by Bekessy et al. [4] is unique in the tropical-disease literature.

5. EVENT RATES

For a two-state homogeneous Markov chain we present formulas for the rates ρ_{12} and r_{12} of transition from state 1 to state 2 (conversion) and rates ρ_{21} and r_{21} of transition from state 2 to state 1 (recovery).

Let $w(t)$, $t \geqslant 0$, be a realization of a two-state continuous-time Markov chain, and $\nu_{ij}(s,t;w)$ be the number of transitions from state i to state j by $w(\cdot)$ in the time interval $[s,t)$. It is well known that

$$P(\nu_{ij}(s,s+h;w) = 1 \mid w(s) = i) = q_i h + o(h)$$

and

$$P(\nu_{ij}(s,s+h;w) > 1 \mid w(s) = i) = o(h), \qquad i = 1,2,$$

where q_i, $i=1,2$, are the parameters in the intensity matrix Q in (3.3) and (3.4). Thus

$$\rho_{ij} = \lim_{h \downarrow 0} h^{-1} E\big(\nu_{ij}(s, s+h; w) \big| w(s) = i\big) = q_i, \qquad i = 1, 2.$$

The intensity q_i may be interpreted as the expected number of transitions from state i to state j per unit time at time s per person at risk of such a transition—that is, in state i, at time s. For time-homogeneous chains, of course, q_i does not depend on s. The seasonal variation in these rates in the Garki population is indicated by comparison of q_i across successive two-wave panels for a given age class (Table 5).

In terms of the counting process $\nu_{ij}(s,t;\cdot)$, $0 \leqslant s < t$, the rates r_{ij} are defined as

$$r_{ij}(s) = \lim_{t \downarrow s} \frac{E\nu_{ij}(s,t)}{t-s}. \tag{5.1}$$

The expectation is calculated over the sample space of all paths $w(t) \in \{1,2\}$ for all $t \geqslant 0$. The sample space is equipped with the σ-algebras of subsets $F_t = \sigma(w(s),\ s \leqslant t),\ t \geqslant 0$, using the probability measure P on $\{F_t,\ t \geqslant 0\}$ associated with a homogeneous Markov chain. Goodman and Johansen [13] give a nice construction of such measures P.

Let $D_i = \{w : w(0) = i\}$, $i = 1,2$, and for $i \neq j$ let

$$N_{ij}(t,w) = \begin{cases} \nu_{ij}(0,t;w) & \text{if} \quad w \in D_i, \\ 0 & \text{otherwise,} \end{cases}$$

$$N_{jj}(t,w) = \begin{cases} \nu_{jj}(0,t;w) & \text{if} \quad w \in D_j, \\ 0 & \text{otherwise.} \end{cases}$$

Then $E\nu_{ij}(s,t) = E\nu_{ij}(0,t) - E\nu_{ij}(0,s)$ and

$$\begin{aligned} E_w \nu_{ij}(0,t) &= E_{D_i}\big(E_w(\nu_{ij}(0,t) | D_i)\big) \\ &= \int_{D_1} \nu_{ij}(0,t;w)\, dP(w) + \int_{D_2} \nu_{ij}(0,t;w)\, dP(w). \end{aligned}$$

In terms of the random functions $N_{ij}(t,w)$, we have, for $i \neq j$,

$$E\nu_{ij}(0,t) = EN_{ij}(0,t)P(D_i) + EN_{jj}(0,t)P(D_j). \tag{5.2}$$

The conditional expectations $EN_{ij}(0,t) = \gamma_{ij}(t)$ satisfy the system of integral equations

$$\gamma_{ij}(t) = G_{ij}(t) + \int_0^t G_{ij}(t-x)\, d\gamma_{jj}(x), \qquad i \neq j, \tag{5.3}$$

where $G_{ij}(t)=1-e^{-q_i t}$ and $\gamma_{jj}(t)=t/l_{jj}-1+l_{jj}^{-1}[p_{jj}(0,t)l_{jj}+p_{ji}(0,t)l_{ij}]$. Here l_{jj} is the mean recurrence time to state j, l_{ij} is the mean waiting time in state i, and $p_{ij}(0,t)=(e^{tQ})_{ij}$. Barlow and Proschan [3, pp. 136–137] give a detailed derivation. Since the waiting times in state i are exponentially distributed with parameter q_i, we have $l_{jj}=1/q_1+1/q_2$. $j=1,2$, and $l_{ij}=1/q_i$, $i=1,2$, $j\neq i$. Thus

$$l_{jj}(t)=\frac{tq_1q_2}{q_1+q_2}-q_1q_2(q_1+q_2)^{-2}(1-e^{-(q_1+q_2)t}), \qquad j=1,2, \tag{5.4}$$

and, integrating (5.3), we obtain for $i\neq j$

$$\gamma_{ij}(t)=\frac{tq_1q_2}{q_1+q_2}+q_i^2(q_1+q_2)^{-2}(1-e^{-(q_1+q_2)t}). \tag{5.5}$$

Finally, substituting (5.4) and (5.5) in (5.2), we obtain for $i\neq j$

$$\begin{aligned} r_{ij}(s)&=\lim_{t\downarrow s}\frac{Ev_{ij}(0,t)-Ev_{ij}(0,s)}{t-s}\\ &=\frac{q_1q_2}{q_1+q_2}+\frac{e^{-(q_1+q_2)s}}{q_1+q_2}\left[q_i^2P(D_i)-q_iq_jP(D_j)\right]. \end{aligned} \tag{5.6}$$

To estimate these event rates from two-wave panel data we use (3.12) for $\hat{q}_i$ and the maximum-likelihood estimators

$$\hat{P}(D_i)=\frac{n_{i+}}{n_{1+}+n_{2+}}, \qquad i=1,2. \tag{5.7}$$

Equations (5.6) and (5.7) are used to estimate conversion and recovery rates from *P. falciparum* parasitemia in the WHO malaria survey (Table 6, Sec. 6). For that table, we assume that the earlier survey in each pair of surveys corresponds to time $t=0$. This is clearly an unrealistic assumption; however, the two-wave panel observations do not allow us to estimate the time between process initiation and the first survey.

Indeed, the formal acceptance of homogeneous Markov chains in the embeddability test of Sec. 3.2 and the estimation of $r_{ij}(s)$ and q_i are as far as one can go in characterizing event rates using 2-wave panel data and the simplest stochastic model consistent with those data.

From (5.6), $\lim_{s\to\infty} r_{ij}(s)=r_{ij}(\infty)=q_1q_2/(q_1+q_2)$ is the equilibrium rate per unit time per individual for transitions from state i to state $j\neq i$. This result may also be derived very simply from the classical epidemiological steady-state relation: prevalence rate (per individual) = incidence rate (per individual per unit time) × mean duration of infection (time). If state 2 is the state of being positive, then in this continuous-time Markov chain the prevalence rate (or "expected equilibrium parasite rate," in the language of

Bekessy et al. [4]) is well known to be $q_1/(q_1+q_2)$, and the mean duration of infection is $1/q_2$. The quotient $q_1q_2/(q_1+q_2)=r_{12}(\infty)$ is therefore the incidence rate. Since a steady state is assumed, $r_{12}(\infty)=r_{21}(\infty)$.

6. RESULTS FROM GARKI BASELINE SURVEYS

Within each age class, recovery rates per positive individual exhibit seasonal variation. The minimum occurs in the early part of the wet season. A maximum occurs toward the middle of the dry season. The seasonal patterns are parallel across age classes, and the recovery rates increase with increasing age (Table 5). The only exception to this pattern is the infants (age < 1), who have higher conversion and recovery rates than persons aged 1–4. The increase in level of $\hat{q}_2$ with age can be attributed to a corresponding increase in immunity. As Bekessy et al. [4] point out, the recovery rate may decline from the infants to age group 1–4 because of the loss of maternal immunity [6] or because of superinfection, which has more of an opportunity to occur in persons at least one year old.

The conversion rates $\hat{q}_1$ are also roughly parallel across age classes. However, as indicated in Table 5, age class 5–8 has a higher rate of conversion than age class 1–4. Above age 9, these rates tend to decrease with increasing age, as already suggested by the increasing immunity.

The recovery rates per individual in the survey, $\hat{r}_{21}$, exhibit the same seasonal patterns as $\hat{q}_2$ within each age class and across age classes when standardized to a common base. For example, Table 8 reflects the same patterns under both standardizations as $\hat{q}_i$ for the same age classes in Table 5. The rates $\hat{r}_{ij}(0)$ in Table 6 are not suitable for making comparisons across age classes in a given two-wave panel survey or across surveys for given age classes, because the rates depend on the initial distributions (5.7). Among the innumerable initial distributions which could be used for standardizations, we apply three to the data from the first two two-wave panel surveys in Table 3, for all 7 age classes (Table 7). The standardizing initial distributions are the initial counts of the age class with the lowest proportion negative, the initial counts of the age class with the highest proportion negative, and the initial distribution between positive and negative of the entire population counting each individual equally. To compare time series of rates for age classes indexed by a variable a, let $n_{i+}(a,s)$ be the initial counts for survey s to $s+1$ in age class a, and let $p_i(a)=[\Sigma_{s=3}^{7}n_{i+}(a,s)]/\Sigma_s\Sigma_i n_{i+}(a,s)$. For the age classes 1–4 and 29–43, we calculated the rates (5.6) using $p_i(a)$ in place of $P(D_i)$ (Table 8).

The essential feature of the rates $\hat{r}_{ij}(0)$ and the various standardizations is that they reflect the influence of the relative sizes of the reservoirs of negative and positive persons at the initial survey on the conversion and recovery rates per individual.

TABLE 6

Malaria Conversion and Recovery Rates at Equilibrium and Assuming $t = 0$ at the Earlier of Each Pair of Surveys[a]

Surveys	Age class	Equilibrium $\hat{r}_{ij}(\infty)$	Conversion $\hat{r}_{12}(0)$	Recovery $\hat{r}_{21}(0)$
3–4	<1	1.995	2.222	1.745
4–5		6.729	24.018	2.179
5–6		2.673	9.460	1.842
6–7		1.699	1.619	1.741
7–8		1.098	0.739	2.079
3–4	1–4	1.432	0.747	1.608
4–5		1.664	3.446	1.518
5–6		1.549	1.639	1.541
6–7		1.469	1.344	1.463
7–8		1.001	1.083	0.993
3–4	5–8	2.479	2.054	2.552
4–5		1.926	3.164	1.815
5–6		2.539	1.986	2.618
6–7		2.730	2.127	2.834
7–8		2.041	2.358	2.000
3–4	9–18	3.490	3.028	3.782
4–5		2.628	4.516	2.145
5–6		3.704	2.944	4.031
6–7		3.756	2.594	4.543
7–8		3.067	2.537	3.471
3–4	19–28	3.418	3.572	3.083
4–5		4.595	6.087	3.003
5–6		5.111	5.159	5.053
6–7		3.549	2.514	6.744
7–8		3.182	3.119	3.347
3–4	29–43	4.437	4.656	3.773
4–5		5.919	7.720	3.428
5–6		5.794	5.481	6.378
6–7		3.443	2.712	6.767
7–8		3.625	3.655	3.511
3–4	44+	4.170	4.295	3.744
4–5		7.343	9.709	3.975
5–6		6.014	5.400	7.340
6–7		3.503	2.834	6.864
7–8		4.224	4.366	3.664

[a]Entries are 1000×rates. Rates are calculated using intensities in Table 5, n_{i+} in Table 4, and (5.6), (5.7).

During the wet season, conversion rates are higher than recovery rates in every age class, because of the high mosquito man-biting rate then. How-

TABLE 7

Standardized Malaria Conversion (r_{12}) and Recovery (r_{21}) Rates per 1000 Days[a]

Surveys	Rates	Age Group						
		<1	1–4	5–8	9–18	19–28	29–43	44+
	Standardized by the minimum proportion uninfected							
3–4	$\hat{r}_{12}(0)$	0.405	0.747	1.824	0.960	0.533	0.629	0.576
	$\hat{r}_{21}(0)$	3.752	1.608	2.591	5.092	9.648	15.990	16.348
4–5	$\hat{r}_{12}(0)$	4.386	2.987	3.164	1.752	1.208	1.385	1.697
	$\hat{r}_{21}(0)$	7.346	1.556	1.815	2.852	8.210	12.185	15.383
	Standardized by the maximum proportion uninfected							
3–4	$\hat{r}_{12}(0)$	3.023	5.568	13.601	7.159	3.977	4.693	4.295
	$\hat{r}_{21}(0)$	0.859	0.368	0.593	1.166	2.210	3.662	3.744
4–5	$\hat{r}_{12}(0)$	25.087	17.087	18.097	10.019	6.913	7.922	9.709
	$\hat{r}_{21}(0)$	1.898	0.402	0.469	0.737	2.121	3.149	3.975
	Standardized by the weighted mean proportion uninfected							
3–4	$\hat{r}_{12}(0)$	2.080	3.832	9.360	4.926	2.737	3.230	2.956
	$\hat{r}_{21}(0)$	1.901	0.815	1.313	2.580	4.888	8.102	8.283
4–5	$\hat{r}_{12}(0)$	17.508	11.925	12.630	6.992	4.824	5.529	6.776
	$\hat{r}_{21}(0)$	3.893	0.824	0.962	1.511	4.351	6.457	8.152

[a]In survey 3, the minimum ratio $n_{1+}/n_{2+}=0.119$ occurred in the age group 1–4; in survey 4, the minimum ratio $n_{1+}/n_{2+}=0.157$ occurred in the age group 5–8. In survey 3, the maximum ratio $n_{1+}/n_{2+}=3.888$ occurred in the age group 44+; in survey 4, the maximum ratio $n_{1+}/n_{2+}=3.478$ also occurred in the age group 44+. The weighted mean proportion uninfected in survey 3 was 0.547, and in survey 4 was 0.542.

ever, when the rates are standardized to a minimum proportion negative (Table 7), the conversion rates are higher than the recovery rates only in age groups 1–4 and 5–8. The older persons, who also have greater resistance to malaria and shorter spells of patent parasitemia, then tend to have higher recovery rates than conversion rates if they are members of a population with a high initial percentage positive. On the other hand, with a large reservoir of initially uninfected persons and the same $\hat{q}_i$, $i=1,2$, all age classes have higher (standardized) conversion than recovery rates.

The steady-state event rates $\hat{r}_{ij}(\infty)$ (Table 6) may also be interpreted as classical epidemiological incidence rates. We will refer to an approximate equilibrium condition as one for which $\hat{r}_{12}(0)\approx\hat{r}_{21}(0)\approx\hat{r}_{ij}(\infty)$, $i\neq j$. As indicated in Table 6, approximate equilibrium occurred for age groups 1–4 through 29–43 at surveys 7–8. These surveys occurred near the end of a dry season, where superinfection was minimal. This approximate equilibrium is temporary, since a new generation of mosquitos arises at the onset of the wet season. For the age group 1–4 there is also an approximate equilibrium

TABLE 8

Comparisons Across Surveys of Malaria Conversion (r_{12}) and Recovery (r_{21}) Rates per 1000 Days[a]

	$\hat{r}_{12}(0)$		$\hat{r}_{21}(0)$	
Surveys	1–4	29–43	1–4	29–43
Age classes 1–4 and 29–43 standardized to age class 1–4				
3–4	0.773	0.651	1.601	15.924
4–5	2.429	1.126	1.601	12.544
5–6	1.932	0.982	1.512	14.768
6–7	1.645	0.464	1.450	16.992
7–8	1.225	0.508	0.979	15.212
Age classes 1–4 and 29–43 standardized to age class 29–43				
3–4	5.026	4.236	0.508	5.048
4–5	15.796	7.323	0.508	3.976
5–6	12.565	6.390	0.479	4.681
6–7	10.698	3.016	0.460	5.386
7–8	7.970	3.303	0.310	4.822

[a]As defined in the text, p_1 (age class 1–4) = 0.110, p_1 (age class 29–43) = 0.718.

at surveys 5–6, in the latter part of a wet season. This apparent equilibrium may result from the enormous imbalance in the ratio of positive to negative persons, 484/50, at survey 5. A high man-biting rate might lead to enough conversions among the 50 negative individuals to balance the recoveries among the many positive persons.

7. SOURCES OF ERROR

7.1. *MEASUREMENT ERROR DUE TO DAILY FLUCTUATIONS IN PARASITE DENSITY*

An individual with patent *P. falciparum* parasitemia exhibits large daily oscillations in parasite density. A blood smear taken at a time of low density in that individual could be misclassified as negative. The distribution of phases of these oscillations in a population of infected persons was not recorded in the Garki surveys. Thus we have no direct measurement of the proportion of infected persons who would be incorrectly classified by blood smears.

~~Despite this lack of numerical information, we describe qualitatively the influence of this misclassification error on embeddability tests, estimates of the parameters q_i, $i = 1,2$, and the event rates $r_{ij}(s)$.~~

~~Consider a model of blood-smear measurement in a two-wave panel study where $(w(0), w(\Delta))$ represent the actual infection states at times 0 and~~

PLEASE REPLACE DELETED TEXT WITH REVISED TEXT IN ERRATUM ATTACHED AT END.

Δ and $(Y(0), Y(\Delta))$ represent the observed states of infection. For $k=0,1$, suppose misclassification errors satisfy

$$\text{Prob}(Y(k\Delta)=1 \mid w(k\Delta)=1)=1, \tag{7.1}$$

$$\text{Prob}(Y(k\Delta)=1 \mid w(k\Delta)=2)=\pi, \qquad 0<\pi<1, \tag{7.2}$$

$$\text{Prob}(w(k\Delta)=2 \mid Y(k\Delta)=2)=1. \tag{7.3}$$

Suppose also that misclassification does not alter the underlying dynamics of infection, so that

$$\begin{aligned}&\text{Prob}(w(\Delta)=1 \mid Y(0)=1,\ w(0)=1)\\ &=\text{Prob}(w(\Delta)=1 \mid w(0)=1)\equiv p_{11}=1-p_{12},\end{aligned} \tag{7.4}$$

$$\begin{aligned}&\text{Prob}(w(\Delta)=1 \mid Y(0)=1,\ w(0)=2)\\ &=\text{Prob}(w(\Delta)=1 \mid w(0)=2)\equiv p_{21}=1-p_{22}.\end{aligned} \tag{7.5}$$

The conditional probabilities p_{ij} are assumed only to be associated with some two-state continuous-time stochastic process. A routine calculation shows that the conditional probabilities for the observed process, $\text{Prob}(Y(\Delta)=i \mid Y(0)=i)=p^*_{ii}$, obey

$$p^*_{ii}<p_{ii}. \tag{7.6}$$

Williams and Mallows [29] study more elaborate models of errors in the determination of p_{ii} due to differential nonresponse in two-wave, two-state panels surveys.

The importance of (7.6) for Markov embeddability tests is that the trace of an estimated transition matrix $\hat{P}^*(0,\Delta)$, based on independent observations $(Y_i(0), Y_i(\Delta))$, $1\leqslant i\leqslant N$, can lead to no decision regarding, or rejection of, the null hypothesis H_0: trace $P>1$, when the unobservable matrix $\hat{P}(0,\Delta)$ satisfies trace $\hat{P}(0,\Delta)>1+\delta_1$ in the formal test of Sec. 3. The diagonal entries in the three matrices in Table 3 leading to no decision might very well underestimate p_{ii} because some infected persons with low parasite density at the survey times were misclassified as uninfected.

If $\hat{P}^*(0,\Delta)$ is accepted as embeddable in a homogeneous Markov chain, then the maximum-likelihood estimator $\hat{Q}^*$ or $(1/\Delta)\log \hat{P}^*(0,\Delta)$ is related to $\hat{Q}$, up to terms of order 2 in the estimated measurement error $\hat{\varepsilon}$, via

$$\hat{q}_{ii}=\hat{q}^*_{ii}+\hat{\eta}_{ii}, \tag{7.7}$$

where

$$\hat{\eta}_{ii}=-\frac{\log x}{x+\hat{\varepsilon}-1}\left[\frac{\hat{\varepsilon}(1-\hat{p}^*_{ii})}{x-1}+\frac{\hat{\varepsilon}_i}{x+\hat{\varepsilon}-1}\right]+\frac{\hat{\varepsilon}(1-p^*_{ii}-\hat{\varepsilon}_i)}{x(x+\hat{\varepsilon}-1)}, \tag{7.8}$$

PLEASE REPLACE DELETED TEXT WITH REVISED TEXT IN ERRATUM ATTACHED AT END.

with $x = \operatorname{trace} \hat{P}^* - 1$, $\hat{p}_{ii} = \hat{p}^*_{ii} + \hat{\epsilon}_i$, and $\hat{\epsilon} = \hat{\epsilon}_1 + \hat{\epsilon}_2$. [By (7.6), for large samples, $\hat{\epsilon}_i \approx \epsilon_i > 0$ and $\hat{n}_i > 0$.] Thus, to within sampling variability, the estimated expected durations in each state implied by $\hat{Q}^*$ are longer than those implied by $\hat{Q}$, i.e.

$$\frac{1}{\hat{q}_i} < \frac{1}{\hat{q}^*_i}. \tag{7.9}$$

The error in $\hat{Q}$ in large samples is of smaller order of magnitude than (7.8).

From (7.9), the estimated equilibrium event rates satisfy, for $i \neq j$,

$$\hat{r}^*_{ij}(\infty) = \left[\frac{1}{\hat{q}^*_1} + \frac{1}{\hat{q}^*_2} \right]^{-1} < \hat{r}_{ij}(\infty). \tag{7.10}$$

However, there is no analogous inequality between the nonequilibrium rates $r_{ij}(s)$ and $r^*_{ij}(s)$ valid for all $s < +\infty$. Indeed, $\operatorname{sgn}[r_{ij}(s) - r^*_{ij}(s)]$ depends on the magnitude of $\hat{\epsilon}_i = \hat{p}_{ii} - \hat{p}^*_{ii}$, and on the proportions of individuals in each state at the first wave of observation.

7.2. *MODEL MISSPECIFICATION*

An important source of heterogeneity among individuals in Garki is differential immunity, even within the age classes used in the WHO malaria surveys. Homogeneous Markov chains treat these heterogeneous populations as though they were homogeneous.

To understand the biases which can arise from treating heterogeneous populations as if they were homogeneous, we describe some mixtures of Markov chains which have $\{P : \operatorname{trace} P > 1\}$ as their reachable set of conditional probabilities and compare expected event rates under these models with the same quantities calculated under a homogeneous Markov model.

Let $t = 0$ be the time of initiation of a mixture of continuous-time Markov chains. Then for every intensity-matrix-valued function $Q(t)$, $t \geq 0$, let $p_i^{(Q)}$, $i = 1, 2$, be the probability distribution of type-Q individuals between states at the initial survey, $t = 0$. Let $p_{ij}^{(Q)}$ be transition probabilities arising from (3.1a, b) with intensity-matrix function $Q(t)$, $t \geq 0$. The transition probabilities $\bar{p}_{ij}(0, t)$ for a mixture of such Markov chains are defined by

$$\bar{p}_{ij}(0,t) = \frac{\int_{\mathbf{Q}} p_i^{(Q)} p_{ij}^{(Q)}(0,t)\, d\sigma(Q)}{\int_{\mathbf{Q}} p_i^{(Q)}\, d\sigma(Q)}, \tag{7.11}$$

where $d\sigma$ is a probability measure on the space $\boldsymbol{Q}$ of measurable intensity-matrix-valued functions $Q(t)$, $t \geqslant 0$. We rewrite (7.11) in terms of the measures

$$d\mu_i(Q) = \frac{p_i^{(Q)} d\sigma(Q)}{\int_{\mathbf{Q}} p_i^{(Q)} d\sigma(Q)} \tag{7.12}$$

as

$$\bar{p}_{ij}(0,t) = \int_{\mathbf{Q}} p_{ij}^{(Q)}(0,t)\, d\mu_i(Q). \tag{7.13}$$

The two subclasses below of this family of heterogeneous population models have the same reachable set of conditional probabilities as would arise if $d\sigma$ were a point mass at one intensity-matrix function. These mixtures are indistinguishable from the continuous-time Markov models using only two-wave panel data:

Class (i) contains those mixtures for which $d\mu_1 = d\mu_2$. Here $\operatorname{trace} \bar{P}(0,t) > 1$ for every probability measure $d\mu_1$ on $\boldsymbol{Q}$.

Class (ii) contains those intensity-matrix-valued functions $Q(t)$ such that $p_{ii}^{(Q)}(0,t) > \frac{1}{2}$ for $t \geqslant 0$. Here $\operatorname{trace} \bar{P}(0,t) > 1$ for every pair of probability measures $(d\mu_1, d\mu_2)$ which are nonzero only at such $Q(t)$.

Arbitrary mixtures (7.13) of stochastic matrices P with $\operatorname{trace} P > 1$ need not satisfy $\operatorname{trace} \bar{P} > 1$. For example, let

$$P^{(Q_1)}(0,\Delta) = \begin{pmatrix} 0.7 & 0.3 \\ 0.6 & 0.4 \end{pmatrix}, \qquad P^{(Q_2)}(0,\Delta) = \begin{pmatrix} 0.4 & 0.6 \\ 0.3 & 0.7 \end{pmatrix}.$$

With the measures $d\mu_1(Q_1) = s_1 = 1 - d\mu_1(Q_2)$ and $d\mu_2(Q_1) = s_2 = 1 - d\mu_2(Q_2)$, $0 < s_i < 1$, $i = 1,2$, we have $\operatorname{trace} \bar{P}(0,\Delta) < 1$ whenever $s_1 - s_2 < -\frac{1}{3}$. This example provides another possible explanation of the matrices in Table 3 which led to no decision in the Markov embeddability test, namely, that these data arose from mixtures of homogeneous Markov chains.

To illustrate the biases in event rates due to model misspecification, consider the homogeneous Markov chains with intensity matrices

$$Q = \lambda \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}, \qquad \lambda > 0. \tag{7.14}$$

For each λ, the equilibrium event rates are $r_{ij}(\infty) = \lambda/2$. Let $d\mu_1(\lambda) = d\mu_2(\lambda)$ be any probability measure on $[0, +\infty)$. Then

$$\bar{p}_{ii}(0,\Delta) = \frac{1}{2} \int_0^\infty (1 - e^{-2\lambda\Delta})\, d\mu_1(\lambda), \tag{7.15}$$

and for the mixture, the equilibrium event rates are $\bar{r}_{ij}(\infty)=\frac{1}{2}\int_0^\infty \lambda\, d\mu_1(\lambda)$. If an estimated transition matrix actually arises from a mixture of the form (7.15) but, following an embeddability test, the Markov models based on (7.14) are used to calculate $r_{ij}(\infty)$, then the actual event rate $\bar{r}_{ij}(\infty)$ will always be overestimated. This bias due to model misspecification follows from Jensen's inequality. If $p_{ii}^{(\lambda)}=(e^{Q\Delta})_{ii}$ and Q is given by (7.14), then

$$2\bar{r}_{ij}(\infty)=\int_0^\infty \lambda\, d\mu_1(\lambda)=\frac{1}{2\Delta}\int_0^\infty \log(2p_{ii}^{(\lambda)}-1)\,d\mu_1(\lambda)$$

$$<\frac{1}{2\Delta}\log\left(2\int_0^\infty p_{ii}^{(\lambda)}\,d\mu_1(\lambda)-1\right)=2r_{ij}(\infty). \tag{7.16}$$

For general homogeneous chains the event-rate biases depend strongly on the mixing measures $d\mu_t$. In the Garki study, if heterogeneity arises from differences in immunity, these mixing measures are estimable only from serological data [8].

8. REVIEW AND CONCLUSIONS

The Garki project is the first longitudinal field survey of malaria with enough parasitological measurements to make possible direct estimates of age- and season-specific incidence rates of conversion and recovery. Formal embeddability tests showed that for each of seven age classes a time series of two-wave panel surveys were indistinguishable from samples of time-homogeneous Markov chains. The two estimated parameters of the intensity matrix associated with each two-wave panel were interpreted as conversion and recovery rates per person at risk of conversion or recovery, respectively. The conversion and recovery rates per person surveyed were computed from the intensities.

Thus, for each age class, the two-state (positive, negative) parasitemia histories $\{X(t,w),\ 2\Delta\leqslant t\leqslant 7\Delta,\ \Delta\approx 10 \text{ weeks}\}$, which evolve during the time span of the baseline surveys in Table 3, are modeled by a time series of two-state continuous-time Markov chains $\{Z_k(t,w),\ k\Delta\leqslant t\leqslant (k+1)\Delta,\ 2\leqslant k\leqslant 7\}$. Here $Z_k(t,\cdot)$ is a Markov chain that represents the parasitological histories only during the 10-week interval between surveys k and $k+1$.

Our discussion of embeddability in Sec. 3 serves to clarify the statement of Bekessy et al. [4]: "if trace $P\leqslant 1$ it could be suspected that the model did not fit well: either the process was not Markovian or the parameters were not constant between subsequent observations." If trace $P<1$, then the underlying continuous-time stochastic process must be non-Markovian. If the underlying process is inhomogeneous Markov, with time-varying intensities, then trace $P>1$. Thus the second proposed explanation is incorrect if the nonconstant parameters occur in a Markov chain.

The estimated recovery rates per positive person increase with increasing age, reflecting a corresponding increasing immunity. Within each age class, conversion rates per negative person are high during the wet season and decrease during the dry season. The reverse pattern holds for recovery rates per positive person.

When conversion and recovery rates per person surveyed in the wet season are standardized to a population with a small proportion of initially negative persons, then recovery rates are higher than conversion rates for all persons over age 8. When the same crude rates are standardized to a high proportion of initially negative persons, then all age classes have higher standardized conversion than recovery rates (Table 7). These standardizations indicate the importance of a large reservoir of negative persons early in the wet season if the high mosquito man-biting rate is to lead to more conversions than recoveries, especially among older persons. On the other hand, in a population with a high proportion of persons initially positive, even the force of superinfection during the wet season is not adequate to overcome the high resistance of older persons to parasitemia. Their short spells of infection lead to more recoveries than conversions in such a population.

Our principal methodological innovations are: (1) a sampling theory of an embeddability criterion, (2) a formal test for the embeddability of 2×2 stochastic matrices, estimated from two-wave panel data, in the class of inhomogeneous continuous-time Markov chains, (3) the use of two-wave panel data to estimate incidence rates per person surveyed within the class of time-homogeneous Markov chains, (4) a qualitative description of biases in estimated rates which can arise from the misclassification of infected persons, and (5) an illustration of biases due to model misspecification that can arise in estimating event rates.

Rates per person at risk were previously estimated by Bekessy et al. [4]. We also present some revised variability assessments for these estimates.

The use of embeddability criteria to decide whether data sampled in discrete time from a continuous-time, discrete-state process are consistent with a proposed class of stochastic models is a fundamental feature of our investigation. This use of embeddability tests is largely undeveloped in the literature on stochastic modeling. In a $(K+1)$-wave panel survey, data on a continuous-time process are of the form

$\{X(k\Delta, i) = \text{state of individual } i \text{ at time } k\Delta,\ k = 0, 1, 2, \ldots, K,\ 1 \leqslant i \leqslant N,$
$N = \text{number of individuals present for } K+1 \text{ waves}\}.$

Here the natural extension of our strategy is to ask whether the estimated multinomial sequence frequencies

$$p_{i_0,\ldots,i_K} = P(X(0) = i_0, X(\Delta) = i_1, \ldots, X(K\Delta) = i_K) \tag{8.1}$$

could have been generated by a proposed class of models. Here $(i_0,\ldots,i_K)$ is a sequence of states—e.g. positive, negative—in which an individual is observed at times $t=0,\Delta,\ldots,K\Delta$, respectively. For example, if the sequence frequencies (8.1) arise from a two-state inhomogeneous continuous-time Markov chain, it must be possible to write them as

$$p_{i_0,\ldots,i_K}=\bar{p}_{i_0}\prod_{j=1}^{K} m_{i_{j-1},i_j}, \tag{8.2}$$

where $(m_{i_{j-1},i_j})=M((j-1)\Delta,j\Delta)$, $1\leqslant j\leqslant K$, are 2×2 stochastic matrices such that

$$\operatorname{trace} M((j-1)\Delta,j\Delta)>1, \tag{8.3}$$

and $\bar{p}_{i_0}$ is the initial distribution of persons among states. The entries m_{i_{j-1},i_j} are interpreted as

$$m_{i_{j-1},i_j}=P(X(j\Delta)=i_j|X((j-1)\Delta)=i_{j-1}). \tag{8.4}$$

The basic point in testing to see if sequence frequencies (8.1) can be represented by expressions such as (8.2) and (8.3) is that *a priori* we need only a sampling theory for the data. Specification of the dynamics follows tests of hypotheses such as embeddability in the class of inhomogeneous continuous-time Markov chains. This kind of strategy is discussed for fitting discrete-time Markov chains to gap-free data by Anderson and Goodman [1]. However, papers on continuous-time modeling [5, 15, 14] estimate parameters and test hypotheses within an *assumed* class of stochastic processes. The fundamental question of whether any model in the proposed class can generate estimated sequence frequencies, such as (8.1), is bypassed entirely.

The algebraic characterizations—analogous to (8.2) and (8.3)—necessary for testing the embeddability of sequence frequencies are largely undeveloped, even for close relatives of the Markov processes such as restricted families of semi-Markov processes and nondegenerate mixtures of such processes. Developing such characterizations represents a major mathematical challenge with an immediate practical payoff. In particular, the unpublished baseline parasitemia data from Garki include sequence frequencies (8.1) with $K=7$. A natural next step would be to test these data for consistency with (8.2) and (8.3), and, if necessary, carry out analogous tests for non-Markovian processes. As we indicated in (4.1), a non-Markovian representation of some of the sequence frequencies in the intervention and follow-up surveys in Garki will be essential.

Special thanks are due Louis Molineaux, World Health Organization, Geneva, and Klaus Dietz, formerly of WHO, for drawing our attention to the Garki malaria surveys and providing very helpful criticism and strong encouragement. Dr. Molineaux also made available many WHO documents of the Garki project, additional unpublished data, and portions of the monograph by Molineaux and Gramiccia (1979) in advance of publication. Helpful comments on an earlier draft were generously contributed by Stephen Fienberg, Niels Keiding, Stanley Wasserman, Harrison White, and two anonymous referees. We thank Thomas Mueller for deriving the explicit formulas needed to calculate (3.13) and for the numerical computations reported in Tables 1–5. Anne Whittaker prepared this typescript. Partial financial support was provided by the World Health Organization and the U.S. National Science Foundation through grants SOC76-17706 to the Center for the Social Sciences, Columbia University for B.S. and DEB74-13276 for J.E.C.

REFERENCES

1 T. W. Anderson and L. A. Goodman, Statistical inference about Markov chains, *Ann. Math. Stat.* 28:89–190 (1957).

2 N. Bailey, *The Mathematical Theory of Infectious Diseases*, 2nd ed., Macmillan, New York, 1975.

3 R. E. Barlow and F. Proschan, *Mathematical Theory of Reliability*, Wiley, New York, 1965.

4 A. Bekessy, L. Molineaux, and J. Storey, The estimation of incidence and recovery rates of *Plasmodium falciparum* parasitemia, from longitudinal data, *Bull. WHO* 54:685–693 (1976).

5 P. Billingsley, *Statistical Inference for Markov Processes*, Univ. of Chicago Press, Chicago, 1961.

6 S. Cohen, Immunity to malaria, *Proc. Royal Soc. London Ser. B* 203:323–334 (1979).

7 R. Cornille-Brögger and H. M. Mathews, Field application to malaria studies of the passive haemagglutination (PHA) test with lyophilized cells, *Bull. WHO* 52:39–42 (1975).

8 R. Cornille-Brögger, H. M. Mathews, J. Storey, T. S. Ashkar, S. Brögger and L. Molineaux, Changing patterns in the humoral immune response to malaria before, during, and after the application of control measures: A longitudinal study in the West African savanna, *Bull. WHO* 56:579–600 (1978).

9 J. R. Cuthbert, On uniqueness of the logarithm for Markov semi-groups, *J. London Math. Soc.* 4:623–630 (1972).

10 K. Dietz, L. Molineaux, and A. Thomas, A malaria model tested in the African savannah, *Bull. WHO* 50:347–357 (1974).

11 K. Dietz, Models for parasitic disease control, in *Bulletin of the International Statistical Institute, Proceedings of 40th Session of I.S.I.*, Warsaw, 1975 Book 1, Invited Papers, pp. 531–544; discussion, pp. 569–570.

12 G. S. Goodman, An intrinsic time for non-stationary finite Markov chains, *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* 16:165–180 (1970).

13 G. S. Goodman and S. Johansen, Kolmogorov's differential equations for non-stationary countable state Markov processes with uniformly continuous transition probabilities, *Proc. Cambridge Philos. Soc.* 73:119–138 (1973).
14 I. Iosifescu and P. Tautu, *Stochastic Processes*, Vol. 2, Springer, New York, 1973.
15 N. Keiding, Maximum likelihood estimation in the birth and death process, *Ann. Stat.* 3:363–372 (1975).
16 J. F. C. Kingman, The imbedding problem for finite Markov chains, *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* 1:14–24 (1962).
17 E. Lehman, A theory of some multiple decision problems, II, *Ann. Math. Statist.*, 28:547–572 (1957).
18 A. Lotka, Contribution to the analysis of malaria epidemiology, I–V. *Amer. J. Hygiene* 3, Supplement, 1–121 (1923).
19 G. Macdonald, *The Epidemiology and Control of Malaria*, Oxford U.P., London, 1957.
20 E. Martini, *Berechnungen und Beobachtungen zur Epidemiologie und Bekämpfung der Malaria*, Gente, Hamburg, 1921.
21 L. Molineaux and G. Gramiccia, *The Garki Project: Research on the Epidemiology and Control of Malaria in the Sudan Savannah of West Africa*, WHO, Geneva, to appear.
22 L. Molineaux, G. R. Shidrawi, J. L. Clarke, R. Boulzaquet, T. Ashkar, and K. Dietz, The impact of propoxur on *Anopheles gambiae* S.L. and some other anopheline populations, and its relationship with some pre-spraying variables, *Bull. WHO* 43:379–390 (1976).
23 J. A. Najera, A critical review of the field application of a mathematical model of malarial eradication, *Bull. WHO* 50:449–457 (1974).
24 R. Ross, *The Prevention of Malaria*, 2nd ed., Murray, London, 1911.
25 R. Ross, *Studies on Malaria*, Murray, London, 1928.
26 P. F. Russell, L. S. West, R. D. Manwell, and G. Macdonald, *Practical Malariology*, Oxford U.P., London, 1963.
27 M. Sheps and J. Menken, *Mathematical Models of Conception and Birth*, Univ. of Chicago Press, Chicago, 1973.
28 B. Singer and S. Spilerman, The representation of social processes by Markov models, *Amer. J. Soc.* 82:1–54 (1976).
29 W. H. Williams and C. L. Mallows, Systematic biases in panel surveys due to differential nonresponse, *J. Amer. Statist. Assoc.* 65:1338–1349 (1970).
30 M. Young, Malaria, in *A Manual of Tropical Medicine* (G. W. Hunter, W. W. Frye, and J. C. Swartzwelder, Eds.), W. B. Saunders, Philadelphia, 1966, pp. 316–359.

Erratum

B. Singer and J. E. Cohen, Estimating Malaria Incidence and Recovery Rates from Panel Surveys, *Math. Biosci.* 49:273–305 (1980)

Dr. Jerry Nedelman kindly informed us that the inequality $p_{ii}^* < p_{ii}$, (7.6) on p. 298 in our original text, is only correct under all circumstances for $i = 2$. This has important implications for the interpretation of the influence of measurement error on embeddability tests. Thus all the text starting with the last three words on line 13, p. 298 through the end of Section 7.1 on p. 299 should be replaced with the following two paragraphs.

Subject to (7.1)–(7.5), the conditional probabilities for the observed process, $\text{Prob}(Y(\Delta) = i \mid Y(0) = i) = p_{ii}^*$, are related to π, $z = \text{Prob}(Y(0) = 1)$, and p_{ii} for $i = 1, 2$ via the formulas

$$p_{11}^* = p_{11}\left(1 - \frac{\pi}{z}\right) + \frac{\pi}{z}\left[1 - (1 - z)p_{22}\right] \tag{7.6}$$

and

$$p_{22}^* = (1 - \pi)p_{22}. \tag{7.7}$$

Observe that $\pi/z < 1$ is a consequence of the relation $\text{Prob}(w(0) = 1) = (z - \pi)(1 - \pi)^{-1}$, where $w(0) =$ actual infection status at time 0. Thus if $T =$ trace $P(0, \Delta)$ and $T^* =$ trace $P^*(0, \Delta)$, we have

$$T^* = T + \frac{\pi}{z}(1 - T). \tag{7.8}$$

The importance of (7.8) for Markov embeddability tests is that it shows that the effect of measurement error is always to move the observed trace, T^*, closer to a zone of no decision than the actual trace, T. Furthermore, if $T > 1 + \delta_1$, then T^* can at worst lead to an interpretation of no decision concerning H_0: trace $P > 1$. The opposite conclusion, namely $T < 1 - \delta_2$, cannot arise as a result of measurement error, since $T^* \geqslant 1$ whenever $T \geqslant 1$. Similarly, a reversal of conclusion via measurement error is impossible when $T < 1 - \delta_2$. A modification of the formal embeddability test which takes account of π and z is to replace δ_i by $\delta_i^* = (1 - \pi/z)\delta_i$ for $i = 1, 2$. This guarantees that

$$1 + \delta_1^* < T^* \quad \text{iff} \quad 1 + \delta_1 < T$$

52 Vanderbilt Ave., New York, NY 10017 0025-5564/82/09151+02$02.75

and

$$1-\delta_2^* > T^* \quad \text{iff} \quad 1-\delta_2 > T.$$

If $\hat{P}^*(0,\Delta)$ is accepted as embeddable in a homogeneous Markov chain, then the maximum likelihood estimator $\hat{Q}^* = (1/\Delta)\log \hat{P}^*(0,\Delta)$ is related to $\hat{Q}$, up to terms of order 2 in the estimated measurement error $\hat{\varepsilon} = \hat{\varepsilon}_1 + \hat{\varepsilon}_2$, via

$$\hat{q}_i = \hat{q}_i^* - \frac{\log x}{x-1}\left(\hat{\varepsilon}_i + \frac{\hat{\varepsilon}}{x-1}(1-p_{ii}^*)\right) + \frac{\hat{\varepsilon}}{x}\frac{1-p_{ii}^*}{x-1},$$

where $x = \text{trace}\,\hat{P}^* - 1$ and $\hat{\varepsilon}_i = \hat{p}_{ii} - \hat{p}_{ii}^*$.