Epidemiol. Infect. (2003), **131**, 757–770. © 2003 Cambridge University Press
DOI: 10.1017/S0950268803008574 Printed in the United Kingdom

Seroepidemiology of hepatitis B virus in Addis Ababa, Ethiopia: transmission patterns and vaccine control

A. ABEBE[1], D. J. NOKES[2*], A. DEJENE[3], F. ENQUSELASSIE[2,4], T. MESSELE[5] AND F. T. CUTTS[6]

[1] *Virology and Rickettsiology Research Team, Ethiopian Health and Nutrition Research Institute* (*EHNRI*), *Addis Ababa, PO Box 1242, Ethiopia*
[2] *Department of Biological Sciences, University of Warwick, Coventry CV4 7AL, UK*
[3] *Biostatistics and Health Service Research Team, Ethiopian Health and Nutrition Research Institute, PO Box 1242, Addis Ababa, Ethiopia*
[4] *Department of Community Health, Faculty of Medicine, University of Addis Ababa, PO Box 1176, Addis Ababa, Ethiopia*
[5] *Immuno-Haematology and Pathology Research Team, Ethiopian Health and Nutrition Research Institute, PO Box 1242, Addis Ababa, Ethiopia*
[6] *Department of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1N 7HT, UK*

(*Accepted 13 February 2003*)

SUMMARY

A community-based seroepidemiological survey of Addis Ababa, Ethiopia was conducted in 1994 to inform on the transmission dynamics and control of hepatitis B virus (HBV) infection. Venous blood from 4736 individuals under 50 years of age from 1262 households, selected using stratified cluster-sampling, was screened for HBV markers using commercial ELISAs. HBsAg prevalence was 7% (95% CI 6–8), higher in males (9%; 7–10) than females (5%; 4–6). HBeAg prevalence in HBsAg positives was 23% (18–29), and less than 1% of women of childbearing age were HBeAg positive. Overall HBV seroprevalence (any marker), rose steadily with age to over 70% in 40–49 year olds, indicating significant childhood and adult transmission. Estimated instantaneous incidence was 3–4/100 susceptibles/year, higher in males than females in 0–4 year olds, and peaking in early childhood and young adults. The age at which 50% had evidence of infection was around 20 years, and the herd immunity threshold is approximated at 63–77%. Addis Ababa is of intermediate-high HBV endemicity, with negligible perinatal transmission. Our main findings are the identification of a significant difference between males and females in the age-acquisition of HBV infection, and marked differences between age groups in HBV incidence rates. These results should target future research studies of underlying risk factors. Furthermore, we generate a crude estimate of the level of coverage of HBV vaccine that would be required to eliminate the virus from the study population.

INTRODUCTION

Cross-sectional community studies of serological markers of HBV infection have an important role in identifying population endemicity, possible routes of transmission and associated risk factors [1–9]. Such data help in the development of appropriate control measures, for example, estimating the optimal age for vaccine delivery, the level of vaccine coverage required for elimination, and possible

* Author for correspondence: Welcome Trust Research Laboratories, Kenya Units, PO Box 230, Kilifi, Kenya.

high-risk target groups [10, 11]. They support predictive mathematical models by which to explore the merits of different control options [11–13], to assess cost-effectiveness of intervention programmes [14], and to predict the risk of evolution of escape mutants [15].

Expansion of the HBV survey database is most pressing in the developing world, where the burden of HBV associated liver disease, and the need for vaccine intervention, is highest [16]. In Africa, in particular, uncertainty remains in the relative importance of different routes of HBV transmission and in the underlying risk factors for infection especially for childhood horizontal transmission [16]. Whilst there are numerous serological surveys reported in the literature, few have been designed to obtain data representative of the socio-behavioural and demographic structure of the community. Greater attention needs to focus upon adequate sample size and sampling design to obtain precise and unbiased incidence and prevalence estimates. Such data can be used to provide realistic estimates of the level of coverage needed to control infection. Additionally, finely age-stratified data support investigation of heterogeneity, by age and gender in incidence rates; providing insight into possible risk factors for, and routes of, transmission and the potential impact of community vaccination [10].

There is a body of literature on hepatitis B in Ethiopia relating its public health importance [17–19], and giving information on the pattern of HBV markers in various geographical locations, age and social groups, and risk categories [9, 20–28]. Our study builds upon this earlier work, providing significantly more detailed and representative data on HBV marker seroprevalence for the capital city of Addis Ababa. Through this it has been possible to characterize the transmission patterns of HBV and define vaccine coverage required for control for an urban developing country population.

METHODS

Emphasis is given to details not presented in previous related publications [29, 30]. Ethical approval was obtained from authorities in Ethiopia (Ethical Committee of the Ethiopian Health and Nutrition Research Institute) and the United Kingdom (St Mary's Research Ethics Committee, London University).

Survey design and implementation

Addis Ababa, the capital city of Ethiopia situated on the main plateau (altitude approx. 2000 m), has a population of 2·1 m and a growth rate of 4%/annum (1994 census). Demographic change in recent decades has arisen from decreasing fertility (the crude birth rate, CBR, has fallen from 35·5/yr in 1967 to 15·5/yr in 1994), and high migration [31, 32]. Administratively, the city is organised into 6 zones, comprising 28 Waredas, sub-divided into 305 urban dwellers associations or Kebeles (large housing estates) of roughly 500–3000 households (see Fig. 1). All households are required to register with the Kebele office. A multi-stage cluster sample design [33] was adopted, with two strata formed by the inner (higher density, $\sim 38\,000/km^2$) and outer (lower density, $\sim 10\,000/km^2$) city (note however that the average number of occupants per household does not differ between strata). The Kebele formed the first stage sampling unit, and the household the second stage. For age categories of single years up to age 4, and 5 yearly from 5–49 years, and assuming a design effect of 2 (in the absence of prior information) [33], a sample size of 150 per age group would generate precision of 7–12% across the expected range of seroprevalence. An estimated 1200 households were needed to reach this sample size. We selected 20 Kebeles in each stratum using probability proportional to estimated size following standard procedures (Fig. 1) [34]. Within each of the Kebeles 35 households (to allow for 15% non-participation) were selected by simple random sampling using the registers held at the Kebele office. In the city periphery some registers required updating due to recent house construction. Blood samples were requested from all children under age 5 years, and adults 15–49 years, and 1 in 2 children of school age (5–14 years), as this age group had the highest representation in the population.

Survey teams carried out household visits over the period May to October of 1994. A household represented all individuals of a single family using the same kitchen and sleeping in the same house, inclusive of helper and visitors at time of survey. One household was selected at random from houses with multiple families. Informed consent was obtained from the head of the household. Household heads were interviewed using a pre-tested questionnaire on indicators of socio-economic status, including years of education of household head, house construction and ownership, number of rooms and beds, utilities

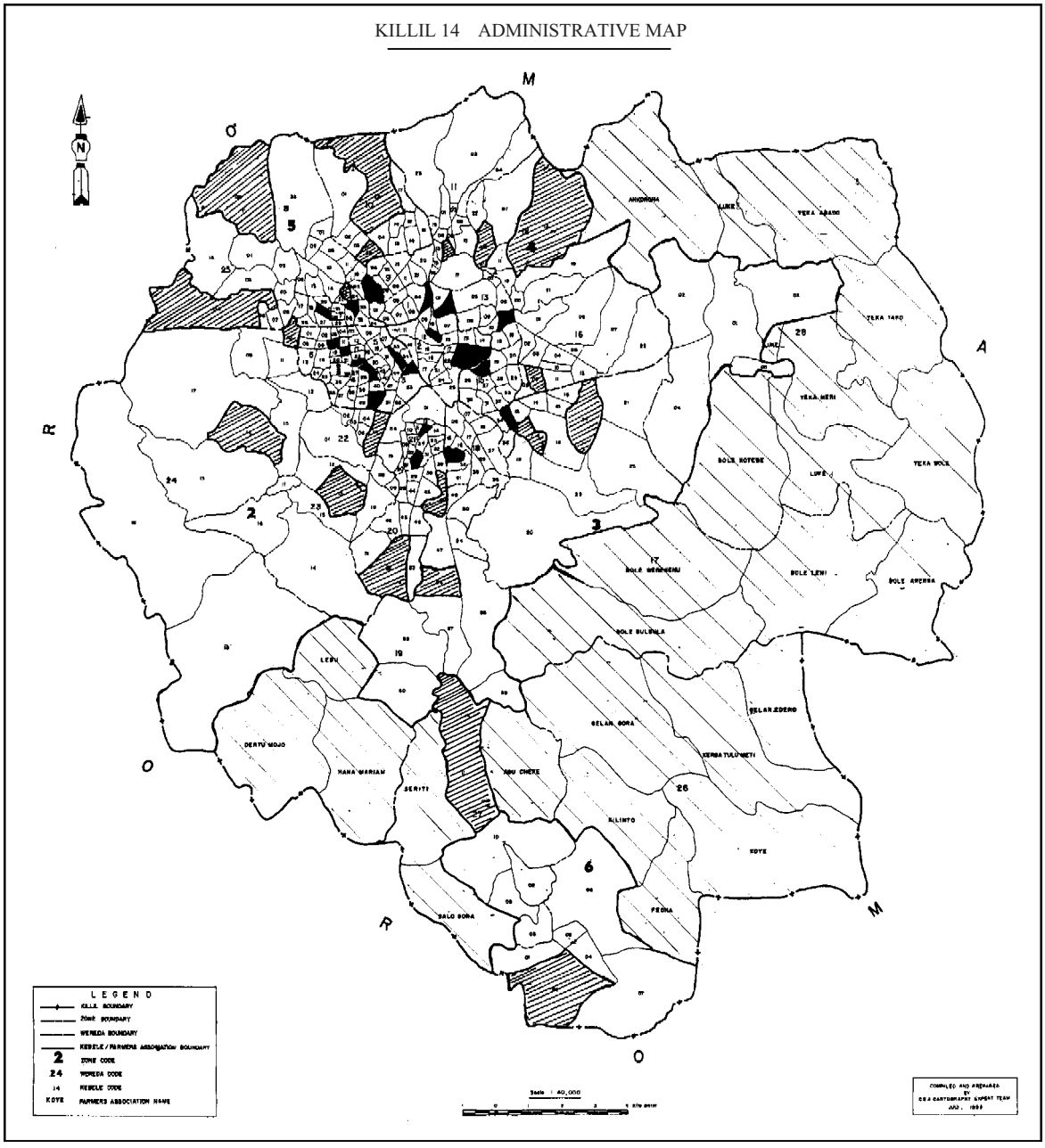

Fig. 1. The administrative structure of Addis Ababa showing the distribution of Kebeles sampled from the inner (black, $n=20$) and outer (dark shading, $n=20$) strata of urban Addis Ababa. Note that Farmer's Associations (denoted by light shading and named) comprising rural Addis Ababa were excluded from the sampling frame.

(water and light) and possession of radio. For each individual in the household, information was requested for age, sex, ethnicity, relationship to household head, years of residence in Addis Ababa, and number of individuals who sleep in the same room. Where possible, 5 ml of venous blood were obtained from family members using sterile disposable needles (21-23G) and Vacutainers (Becton Dickinson, Banbury, England). Samples were transported in cool boxes with ice blocks to the virology laboratory at the EHNRI at the end of each day. Where needed, at least two revisits were made to households to complete interviews and blood sampling.

Laboratory screening methods

Serum collected from centrifuged clotted bloods was stored in screw cap vials marked with a unique

identifier in three aliquots at −20 °C. Primary and confirmatory screening for antibodies to HBV core antigen (anti-HBc) and for HBV surface antigen (HBsAg) of all samples, and quantitative screening for antibodies to hepatitis B surface antigen (anti-HBs) of HBsAg negative samples, were carried out using the Hepanostika range of kits (Hepanostika AHBC Uni-Form; HBsAg Uni-Form II; AHBS New, Organon Teknika, Boxtel, Belgium). HBsAg positive samples were screened for HBeAg (Wellcozyme HBeAg/anti-HBe, Murex, Dartford, England). All kits were used according to the manufacturer's instructions, excepting the anti-HBs assay where samples with an excess of 100 IU/L were not repeat screened at higher dilution.

Data management and statistical analysis

In all data analyses account is taken of the stratified cluster-based study design [33]. Probability weights are applied to each outcome observation, with increased weight given to data from the outer city due to its greater size (inner : outer; 881781 : 1202807) and to data from household clusters (Kebeles) with less than average number of observations (since equal numbers of households were expected from each cluster). Thus observations from cluster i ($i=1$; $n=40$) in stratum j ($j=1, 2$; 1 = inner, 2 = outer) are given the following weighting, W_{ij}

$$W_{ij}=\frac{\sum_{i=1}^{n} c_i}{nc_i}\cdot\frac{2s_j}{\sum_{j=1}^{2} s_j},$$

where the number of households selected in the ith cluster is c_i, and the census population size in the jth stratum is s_j. Furthermore, variance estimation for statistical testing took into account the degree of inter- versus intra-cluster variability for the first stage of cluster sampling, i.e. the Kebele. The analysis thus makes some account for possible lack of independence of observations within clusters of households, though not explicitly between observations within households. Estimates for total and adult population prevalence of HBV markers (15–49 years) are presented standardized for differences in the age distribution of the survey population (those providing a blood sample) and the 1994 census. The weighting for an observation in cluster i, stratum j, age class k ($k=0$–4, …, 45–49) and gender l is $W_{ijkl}=W_{ij},W_{kl}$, with $W_{kl}=y_{kl}/z_{kl}$ being the weighting applied to a serological observation in inverse proportion to the representation of its age class and gender in the survey $z_{kl}=Z_{kl}/Nb$ relative to the census $y_{kl}=Y_{kl}/Nc$. Here Z_{kl} and Nb are the numbers of blood samples in the survey that are of age class k and gender l, and in total, respectively, and Y_{kl} and Nc the number of individuals in the censused population that are in age class k and gender l, and in total, respectively. All statistical procedures were implemented in STATA V6 (Stata Corporation, College Station, Texas, USA) using the 'svy' commands, e.g. 'svytab' and 'svylogit'. STATA svytab command produces asymmetric 95% confidence intervals (CI) for prevalence estimates using a logit transform to ensure they lie between 1 and 0. We use the svylogit command to estimate Odds Ratios (OR), with 95% CI, for comparisons between prevalence in different groups (e.g. males and females).

Estimation of the rate of HBV transmission

Age-dependence in the force of infection (instantaneous per susceptible incidence) for hepatitis B in the Addis Ababa population was investigated using a piece-wise constant (PWC) catalytic infection model [30]. The analysis ignored hepatitis-B related deaths, and assumed that all infected individuals develop a detectable serological response and retain at least one serological marker up to at least age 49 years (the upper age in the survey). We also assumed that HBV transmission has not changed significantly over the last 50 years in the population. Under these assumptions, the change with age in the proportion serologically positive reflects the age acquisition of HBV infection of a birth cohort, and any variation in this acquisition rate reflects age related phenomena and not time-related changes in HBV transmission. It follows that the proportion remaining susceptible (i.e. serologically negative, having no marker of HBV infection) at age a, $x(a)$, is a function of the cumulative exposure to HBV infection from birth over a years (i.e. the cumulative force of infection or cumulative hazard). We modelled this cumulative incidence using a PWC catalytic infection model, in which the force of infection, λ_i, is assumed constant in age class i, but can vary between age classes (i.e. $i=1, m$), hence the predicted proportion susceptible, $x'(a)$ is defined as

$$x'(a)=x'(a_{i-1})\exp[-\lambda_i(a-a_{i-1})], \quad (1)$$

where $a_{(i-1)}\leqslant a\leqslant a_i$ and $\lambda_i\geqslant 0$. This modelled an exponential decay in the proportion susceptible between age a_{i-1} and age a due to an age-class specific force of infection, λ_i. We assumed there is no influence of

maternal specific antibody and that perinatal transmission was negligible [28], i.e. that all individuals are born susceptible. The predicted proportion serologically positive at age a is simply $p' = 1 - x'(a)$. We applied the model to yearly age group serological data recording the number of individuals without any HBV markers, $S(a)$, out of a total tested, $N(a)$ (each adjusted by probability weights to account for the stratified cluster sampling design), where $S(a)/N(a) = x(a)$ is the seronegative prevalence at age a. Force of infection parameter estimates (and 95% CL) were estimated by maximum-likelihood using previously described methods [30, 35]. A best fit model to the data was derived with minimum number of force of infection age classes using step-wise elimination strategy. First we fitted an 8 parameter model (i.e. $i = 1, 8$), with 5 year age classes from 0–4 years up to 25–29 years, and 10 year classes thereafter, estimating the log-likelihood of the model. Age groups observed to have similar λ_i estimates were grouped and the model refitted. Comparison between the two models was then achieved using the likelihood ratio test of the two log-likelihoods (with degrees of freedom equal to the difference in number of parameters between the two models), and the constrained model accepted if there was no significant reduction in goodness of fit. The process was continued and a minimally specified best-fit model defined as that with fewest parameters (age classes) with a statistical fit to the data no worse than for the 8 parameter model.

Summary epidemiological parameters

Following the methods described elsewhere [30, 36, 37] we calculated the average (arithmetic mean) age at infection, A, the basic reproduction number, R_0, and the critical level of (successful) vaccination required for elimination, p_c (also known as the herd immunity threshold, H), for hepatitis B in the population of Addis Ababa.

We defined incidence in age group a_1 to a_2 as

$$\mathrm{I} = \int_{a_1}^{a_2} \lambda(a) X(a)\, \mathrm{d}a, \tag{2}$$

where $X(a)$ is the number of individuals of age a susceptible to hepatitis B (i.e. without any HBV marker), estimated from the predicted proportion susceptible, $x'(a)$ (see eqn (1)) divided by $N(a)$, the number of individuals of age a in urban Addis Ababa, 1994 [31]. It follows that the average age at infection, A is

$$A = \frac{\int_0^\infty a\lambda(a)X(a)\, \mathrm{d}a}{\int_0^\infty \lambda(a)X(a)\, \mathrm{d}a}. \tag{3}$$

Solutions to eqns (2)–(3) were approximated using discrete age intervals of 0·1 years, with an upper age limit for eqn (3) of 99 years [30]. We adopted the expression $R_0 = 1 + B/A$, where B is the reciprocal of the crude birth rate (per 1000 head of population per year), and $p_c = 1 - 1/R_0$. Estimates of R_0, and thus p_c, are crude approximations only, based upon the simplifying assumptions of age-independence in the force of infection, negligible influence of maternally derived immunity in infancy, and simple demographic approximations [30]. A more detailed analysis is beyond the scope of this paper.

RESULTS

A total of 8638 individuals from 1384 households within 40 clusters (see Fig. 1) were registered in the study, from whom 4777 blood samples were obtained. Further details on the characteristics of the surveyed population are found in other publications [29, 30]. Results are presented for 4736 individuals from 1262 households who were aged 0–49 years and had sufficient serum for serological analysis.

Serological results

Of the 4736 samples screened 292 (6·2%) were HBsAg positive and 1734 (36·6%) were anti-HBc positive. Within the group of HBsAg positive samples 13·7% (40) were anti-HBc negative, of which 24 were in children under 15 years old representing 38% of HBsAg positives in this age group. Of the 4444 HBsAg negative sera, 4414 had sufficient residue for anti-HBs screening, of which 1796 (40·7%) had a level of 10 mIU/ml or greater (seropositive). In the data analysis, the 30 samples with insufficient residue were assigned the same status as for the corresponding anti-HBc results (50% positive), since there was good overall agreement (79%) between serological status defined by anti-HBc and anti-HBs. Sufficient serum from 279 of the 292 HBsAg positive samples was available for HBeAg screening, of which 71 (25·4%) were positive. Of samples with anti-HBs levels $\geqslant$50 IU/l, 20% (236/1198) were anti-HBc negative, whereas 65% (390/598) of the samples with anti-HBs levels of 10–49 IU/l were anti-HBc

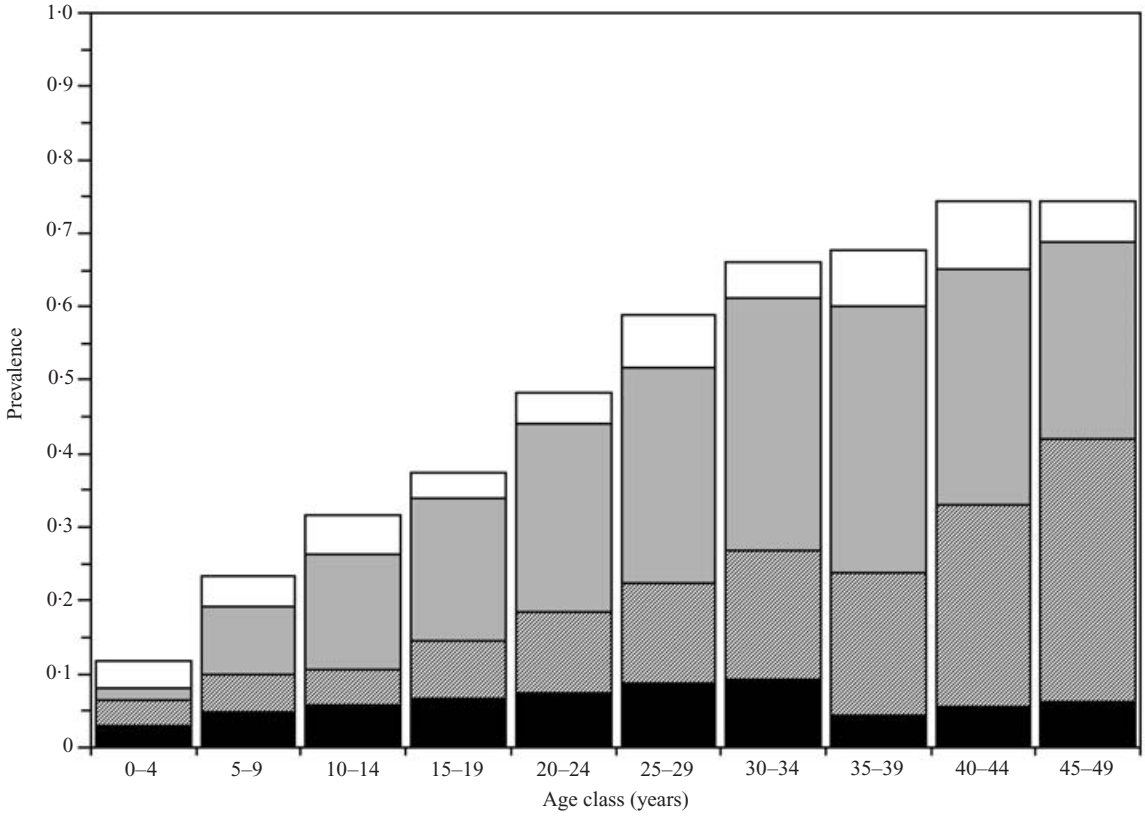

Fig. 2. Age-prevalence of hepatitis B serological markers in Addis Ababa, 1994. The proportions positive for HBsAg, anti-HBc alone, anti-HBc and anti-HBs, and anti-HBs alone, are defined by bars of black, diagonal lines, grey, and open, respectively. Prevalence estimates account for stratified cluster-based study design. Sample sizes for each 5 year age classes were: 0–4, 422; 5–9, 622; 10–14, 796; 15–19, 902; 20–24, 606; 25–29, 397; 30–34, 283; 35–39, 329; 40–44, 191; 45–49, 188; total, 4736.

negative. These 390 samples were considered equivocal (i.e. probable false anti-HBs positives), and in presenting the results the baseline assumption was to assign these anti-HBs equivocal samples as negative, with comparisons made using the alternative assumption.

Prevalence of HBV markers

The prevalence of individuals with at least one marker, standardized to the census population, was 44·8% (95% CI 42·7–47·1), and 55·7% (53·0–58·3) in adults (15–49 years). These estimates become 53·1% (50·5–55·8) and 63·5 (60·8–66·1), respectively, if we assign as positive equivocal anti-HBs samples. Standardized HBsAg prevalence was 6·9% (5·9–8·0), 7·9% (6·8–9·3) for adults (15–49 years).

The prevalence distribution of markers for HBV infection stratified by 5 year age classes is shown in Figure 2. Estimates are weighted to account for stratified cluster sampling. HBsAg prevalence showed a rise from 3·1% (1·7–5·6) in 0–4 year olds to a peak at 9·2% (6·1–13·8) in 30–34 year olds, and subsequent decline to about 5% in 35–49 year olds. The proportion of the population with at least one marker of HBV infection was lowest in the 0–4 year olds at 11·9% (8·5–16·3) and thereafter increased throughout childhood and early adulthood to a plateau in the 40–49 year age group at around 74%. For individuals with a marker of past infection but not current (i.e. excluding HBsAg positives) most had both anti-HBc and anti-HBs markers. The prevalence of individuals with anti-HBs alone remained fairly constant (average of 5%) throughout the age range, whilst the proportion with anti-HBc alone increased with age, particularly in the 40+ years age group, such that in the age group 45–49 years the proportion with anti-HBc alone exceeded that with both anti-HBc and anti-HBs. Assigning equivocal anti-HBs samples as positive led to an increase in overall prevalence of markers, with higher impact in children (standardized childhood seroprevalence is 24·0%

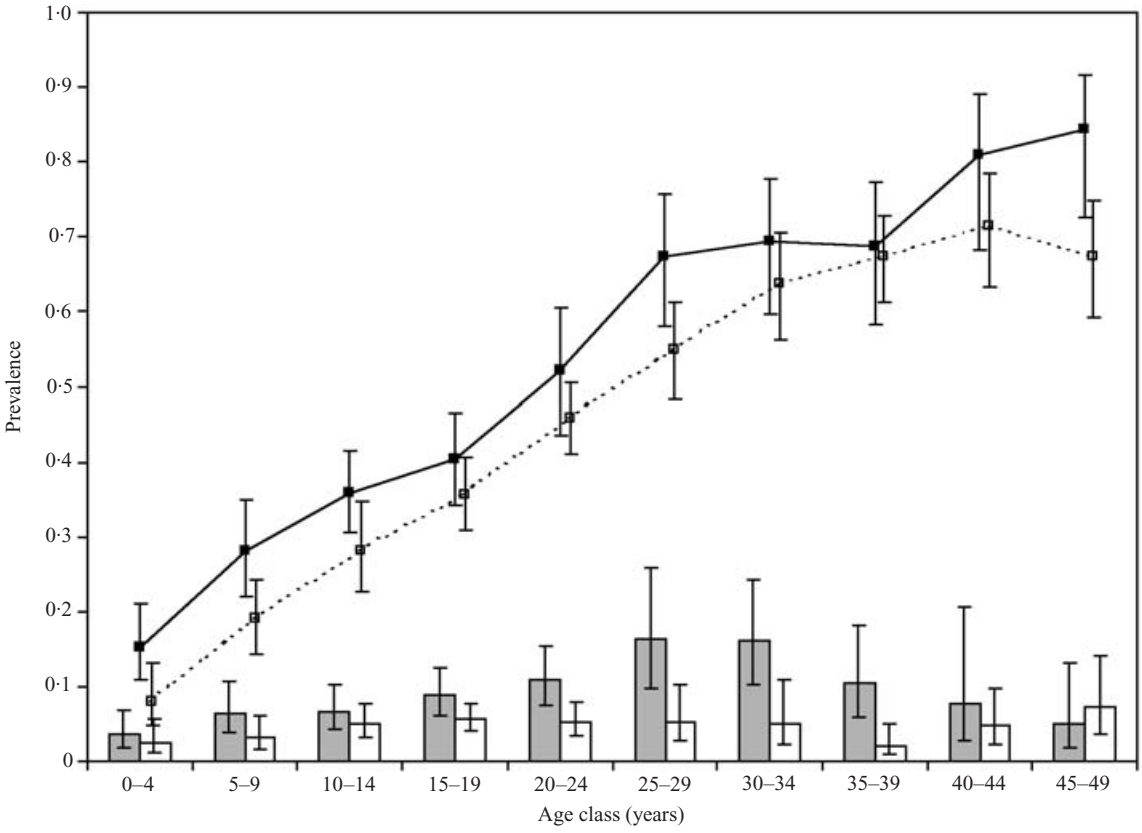

Fig. 3. Age- and sex-stratified prevalence (and 95% CI) of HBsAg (bars) and any marker (lines) in Addis Ababa, 1994. Female prevalence is indicated by open bars and by open markers (dashed line). Prevalence estimates and 95% CI account for stratified cluster-based study design. Sample sizes for each 5 year age class, for males and females, are: 0–4, 225, 197; 5–9, 294, 328; 10–14, 358, 438; 15–19, 326, 576; 20–24, 233, 373; 25–29, 128, 269; 30–34, 108, 175; 35–39, 89, 240; 40–44, 60, 131; 45–49, 77, 111; total, 1898, 2838.

(21·2–27·1) for the baseline assumption and 33·2% (29·6–37·1) assuming equivocal to be positive).

Comparison between males and females in the proportion with the HBsAg marker (bars) or at least one HBV marker (lines) is shown in Figure 3. The prevalence over all ages (0–49 years) of HBsAg in males was 8·6% (7·2–10·2), significantly higher than in females (4·6%; 3·6–5·7) (age-adjusted OR 2·08; 95% CI, 1·59–2·72). The male:female ratio in HBsAg prevalence peaked at between 3:1 and 5:1 in the 25–29 to 35–39 year age classes. The prevalence of individuals with at least one marker was higher for males than females throughout the age range. The overall prevalence of at least one marker for males (44·1%; 39·3–43·7) was significantly higher than for females (41·4%, 42·3–47·0) (age-adjusted OR 1·43; 1·27–1·59).

Seroprevalence (standardized) of the HBeAg marker in HBsAg positive individuals was 23·0% (17·5–29·7), and highly age-dependent, as shown in Figure 4, decreasing from 89·1% (53·1–98·3) in 0–4 year olds to 14·5% (9·5–21·4) in adults (15–49 years). Controlling for age (child *vs.* adult) the prevalence of HBeAg was significantly higher in males (29·4%; 22·1–38·0) than females (19·7%; 13·9–27·2) (adjusted OR 2·05; 1·08–3·15, interaction between age and sex not significant). In women of childbearing age (15–45 years) the prevalence of HBsAg positives was 4·9% (3·8–6·4) in whom the proportion HBeAg positive was 10·4% (456–18·7).

Comparison of marker prevalence in migrants and non-migrants revealed no significant differences. The odds ratios (adjusted for age and sex) for HBsAg and total prevalence in migrants relative to non-migrant were 1·22 (0·89–1·66) and 1·15 (0·98–1·35), respectively.

Incidence of HBV infection

Results of analysis using piece-wise catalytic infection models are in Table 1. Data in single year age groups (0–49 years) were prior weighted to account for

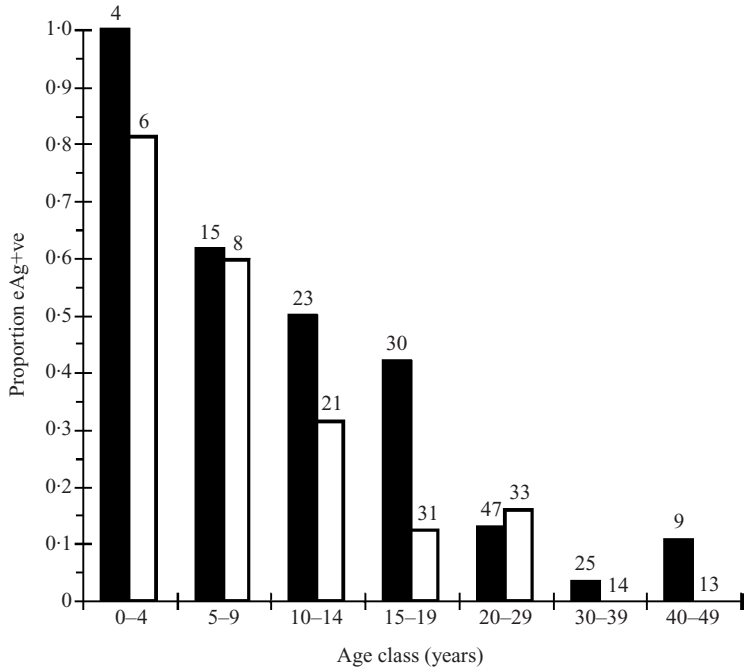

Fig. 4. HBeAg prevalence in HBsAg positives in Addis Ababa, 1994. Prevalence estimates account for stratified cluster-based study design. Sample size for males (filled bars) and females (open bars), respectively, are indicated above each bar, with total for males of 153 and females of 126.

stratified cluster-based design, and analysed in total and by gender, and by assignment of equivocal anti-HBs results as negative or positive. In each case results are presented for the best fit age-dependent model (2 parameters in each case) and the age-independent (constant) model (1 parameter), showing the estimated force of infection (with 95% CI) for each age class, and likelihood estimate for the model fit to the data ($-$log-likelihood). The likelihood ratio test (lr test) comparisons all showed significantly improved fit to the data for the 2 parameter models compared to the 1 parameter model, and a fit to the 2 parameter model that was no worse than that by the 8 parameter model.

Results shown in Table 1, under the assumption that equivocal anti-HBs results are seronegative, indicated an average force of infection (instantaneous per susceptible incidence) for HBV of 3·1%/annum, (i.e. 3·1/100 susceptibles/annum), slightly higher for males and lower for females (constant model). However, age-dependence in transmission is supported statistically indicating peak incidence rate in the total population of 4·4%/annum in 0–4 year olds and 20–29 year olds, and a lower incidence of 2·2%/annum in 5–19 year olds and 30–49 year olds (Table 1). The same age-related pattern was detected for males, with peak incidence of 5·5%/annum, but for females, a higher incidence in the 0–4 year olds was not supported statistically, only in the 20–29 year age group (4·1%/annum). Incidence estimates by gender are graphically presented in Figure 5, together with the fit of the best model to single year age-prevalence data.

By comparison, analysis in which equivocal anti-HBs were assigned as positive lead to an overall increase in force of infection estimates (average for total data 4·0%/annum; 3·9–4·2) with similar but more exaggerated age-dependent patterns for the force of infection, including a significant second peak incidence in females in age group 0–4 years.

Summary epidemiological parameters

Around 50% of individuals had at least one marker by age 20–24 years (Fig. 2). Given the age structure of the population and the age-related forces of infection (Table 1), the modal ages for incidence, I, were 0–4 years and 20–24 years. The arithmetic mean age at infection, A was 17·6 years (males 16·8 years; females 18·7 years). Resultant estimates for R_0 lay between 2·7 and 4·1 (dependent upon the assumed crude birth rate of between 35·5 and 15·5/1000 per annum [30]), from which the critical fraction to vaccinate effectively for elimination is estimated to lie in the range 63–75%.

Table 1. *Force of infection estimates for HBV in Addis Ababa, 1994*

Equivocal assignment*	Group	Model (parameters)	Age classes (years)	FOI estimate (%)	95% CI	−(log-likelihood)	Lr test†
Negative	Total	Constant (1)	0–49	3·1‡	2·9–3·2	2905·06	2*vs.*1 15·60 $P=0·0001$
		Best fit (2)	0–4/20–29	4·4	4·1–4·7	2897·26	2*vs.*8 3·30 $P=0·7704$
			5–19/30–49	2·2	2·0–2·4		
	Males	Constant (1)	0–49	3·6	3·3–3·9	1158·16	2*vs.*1 12·72 $P=0·0004$
		Best fit (2)	0–4/20–29	5·5	4·9–6·1	1151·80	2*vs.*8 4·06 $P=0·6686$
			5–19/30–49	2·3	1·9–2·7		
	Females	Constant (1)	0–49	2·8	2·6–3·0	1727·35	2*vs.*1 5·01 $P=0·0252$
		Best fit (2)	0–19/30–49	2·6	2·4–2·8	1724·84	2*vs.*8 2·74 $P=0·8399$
			20–29	4·1	3·2–5·0		
Positive	Total	Constant (1)	0–49	4·0	3·9–4·2	3030·29	2*vs.*1 52·22 $P<0·0001$
		Best fit (2)	0–4/20–29	6·8	6·4–7·2	3004·17	2*vs.*8 10·81 $P=0·0946$
			5–19/30–49	2·2	1·9–2·5		
	Males	Constant (1)	0–49	4·6	4·3–4·9	1197·51	2*vs.*1 32·38 $P<0·0001$
		Best fit (2)	0–4/20–29	8·2	7·4–8·9	1181·32	2*vs.*8 3·82 $P=0·7007$
			5–19/30–49	2·2	1·7–2·7		
	Females	Constant (1)	0–49	3·7	3·5–3·9	1819·29	2*vs.*1 17·38 $P<0·0001$
		Best fit (2)	5–19/30–49	5·7	5·2–6·2	1810·60	2*vs.*8 7·09 $P=0·3124$
			0–4/20–29	2·4	2·1–2·7		

* Indicates whether equivocal anti-HBs samples were assigned as seronegative or seropositive (see text for details).

† Lr test, likelihood ratio test. 2*vs.*1 and 2*vs.*8 indicate comparisons between models of different numbers of parameters. −log-likelihood for 8 parameter model (equivocals assigned negative): total 2895·61; male 1149·77; female 1723·47. (equivocals assigned positive): total 2998·77; male 1179·41; female 1807·05.

‡ 3·1%/annum is per susceptible incidence, i.e. 3·1/100 susceptibles/annum.

Repeated analysis assigning as seropositive the equivocal low anti-HBs samples lead to no substantial change in these estimates; the predicted values for A, A_{50}, R_0 and H become 16·2 and 20 years, 2·8–4·2, and 64–76%, respectively.

DISCUSSION

Data are presented for an unusually large and detailed serological survey of hepatitis B virus in an urban sub-Saharan African population. The sample was broadly representative of the population of Addis Ababa judged by similarity with the 1994 census in age- and sex-structure [29, 30] and in various social, economic and demographic characteristics (e.g. ethnicity, educational status, migrancy, household size [38]. Estimates of marker prevalence took account of the stratified cluster-based design. Where appropriate prevalence estimates were standardized to the population census to account for age and sex bias in blood sample response rates (adult males being the least willing [29, 30, 38]).

HBV prevalence (past or current infection) was estimated to lie somewhere between 45% and 53% (56–63% in adults aged 15–49 years), with HBsAg prevalence (i.e. current infection) of 7% (in adults 8%). A steady rise in seroprevalence (any marker) throughout childhood was observed, continuing in adulthood, rising to 70–80% in those aged 45–49 years, indicative of significant adult and childhood

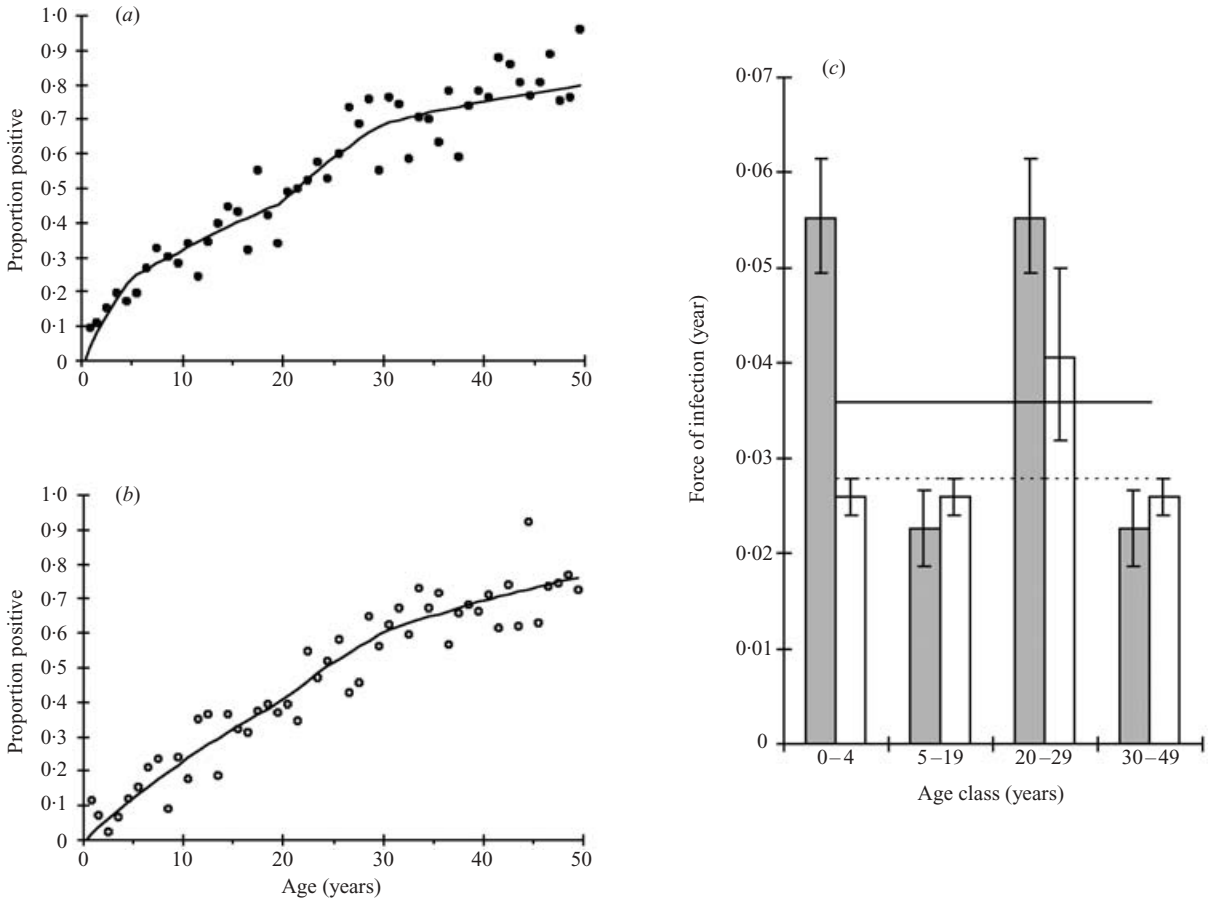

Fig. 5. Age- and sex-related pattern of HBV incidence in Addis Ababa, 1994. (*a*) Shows for males the best fit (2 parameter) piece-wise constant (PWC) catalytic infection model (solid line) to yearly age–prevalence data (i.e. proportions with at least one marker) (markers). (*b*) For females. (*c*) Records the estimated age-specific forces of infection (/susceptible/year), and 95% CI, from the best-fit PWC (males, grey bars; females, open bars). For comparison the age-independent force of infection estimate is shown (males, solid line; females, dotted line).

transmission. Although one quarter of individuals currently infected (HBsAg positive) were positive for a marker of active viral replication and high infectiousness (HBeAg), only about 5 in 1000 women of childbearing age were estimated to be highly infectious, that is, of the 5% HBsAg positive, 11% were HBeAg positive. This suggests the risk of perinatal transmission to be low. Many of these features are in common with surveys from other countries in East Africa, such as Kenya [7, 39], and intermediate endemicity countries in Mediterranean North Africa, such as Tunisia [40, 41]. Significant childhood horizontal transmission, and a HBsAg prevalence in excess of 5–8%, are features ascribed to high endemicity [8, 42]. Nevertheless, within sub-Saharan African countries adult surface antigen prevalence is frequently well in excess of 10%, and the proportion of adults with evidence of HBV infection in excess of 75% [8]. The contrast in HBV transmission rates is particularly acute with West African countries (e.g. The Gambia [11] and Senegal [1, 5]). A useful comparative measure of the magnitude of transmission is the age at which 50% of the population have experienced HBV infection, A_{50} [16]. In Addis Ababa the value of A_{50} is around 20 years of age, compared with less than 10 years in The Gambia and Senegal. These comparisons suggest Addis Ababa to be of intermediate–high HBV endemicity.

Previous serological data from Ethiopia are generally consistent with the findings reported here. A countrywide study published in 1986 [9] recorded HBV prevalence of 40% (76% in adults >15 years), HBsAg prevalence of 6% (12% in adults), a prevalence of HBeAg in HBsAg positives of 26% (9% in adults) and an A_{50} >15 years. Data for Addis Ababa from the same study showed no differences from these averages. A national survey of male military recruits [20] gave an average prevalence of HBsAg of 10·8%

and of any marker of 73·3%, and in male blood donors in NW Ethiopia HBsAg prevalence was reported as 14·4% [24]. A study of HBsAg carrier pregnant women in Addis Ababa [28] provided direct evidence of the low risk of perinatal transmission. Whilst estimates of HBsAg prevalence of 10–14% and seroprevalence of 76–79% for Ethiopian adults [9, 20, 24, 27] appear higher than those observed in this study, the low precision of prevalence estimates in some of the earlier studies, and differences between studies in age representations and sampling criteria, make quantitative comparison difficult. Regional differences in HBV marker prevalence are observed in Ethiopia [9, 20, 24], which might affect prevalence in Addis Ababa in view of the considerable in-migration over the past 20 years. However, we could find no evidence of different marker prevalence between long-term residents and migrants (resident <10 years).

We found evidence for the occurrence of false low positive anti-HBs results. Primary analysis was based on assigning as seronegative anti-HBc negative samples with low anti-HBs levels (10–49 mIU/ml), which is the most likely explanation. Assigning equivocal anti-HBs samples as positive, resulting in higher seroprevalence and increased age-related heterogeneity, was viewed as the less likely alternative. The implications of these differences to the summary epidemiological statistics, such as the herd immunity threshold, are not substantial.

Higher prevalence of HBsAg in males compared to females has been observed in numerous studies in low, intermediate and high endemicity countries, and is apparently ubiquitous. This may be attributable to a higher probability following infection of becoming a carrier in males than in females [43–45]. However, other factors may be important, such as a sex differential in the rate of infection (as observed in this study) or in the duration of carriage, and further investigation would be of interest. The HBsAg male: female prevalence ratio of roughly 2:1 (age adjusted odds ratio of 2·1; 95% CI 1·6–2·7) in this study is consistent with previous work in Ethiopia [9], and in Kenya [7]. The pattern of rapid decline in HBeAg prevalence in HBsAg positives has been observed in both African [7] and East Asian [46] communities. It is interesting to note the significantly higher HBeAg prevalence in male HBsAg positives than females in a ratio of roughly 3:2 (age adjusted odds ratio of 2·1, 95% CI 1·1–3·2) and independent of age. Thus, not only are males more likely to be currently infected but also they are more infectious than females. A study in Cameroon [43] observed higher prevalence of HBV-DNA in males than in females in children, indicative of higher infectivity of males perhaps due to slower clearance of the virus.

Age-dependence in the force of infection has been investigated using age-seroprevalence data based upon the assumption of a steady state in HBV transmission in the population. This assumption is reasonable insofar as studies that precede the current survey are similar in marker prevalence, and, although Addis Ababa is undergoing a demographic transition, there is no evidence to link changes in population density with HBV transmission [16, 47].

A main finding from the present study was the identification of age- and gender-related patterns in seroprevalence (any marker) and consequently in HBV incidence per susceptible (termed the force of infection). The average force of infection was calculated to be 3–4% per year (i.e. 3–4/100 susceptibles per annum). However, there was higher seroprevalence (any marker) in males than females throughout the age profile (Fig. 3), leading to a marked gender difference in the age-acquisition of HBV (male 3·6%/year; female 2·8%/year). This observation is unusual in intermediate or high endemicity countries and only previously reported in Kenya (in adults but not children [7], and in a study of Ghanaian children [6]. The difference in seroprevalence between males and females is established early in life (0–4 year age group) with no obvious increase in divergence with age. This suggests an early sex differential in the risk of acquiring HBV, and the estimated HBV incidence for males was over twice that for females at this age. Further age-heterogeneity was identified in the age group 20–29 years in which both male and female seroprevalence rose more steeply than adjacent age groups, resulting in significantly higher incidence estimates.

Few studies have analysed incidence patterns for HBV in developing countries. A review by Edmunds et al. [16] of highly endemic areas (sub-Saharan Africa and East Asia) revealed a general trend for higher incidence (force of infection) in younger children than older. Age-related patterns of HBV incidence may not be readily discernable by eye from age-seroprevalence data, and require statistical analysis of the form used in this study to be resolved. This may be particularly the case for adults where only a relatively small proportion remain susceptible. Our work is the first, of which we are aware, to have analysed seroprevalence data across both childhood and adult age ranges in

a developing country and also to have identified age-variation in adult incidence.

The patterns of incidence identified in this study probably reflect underlying behavioural and social heterogeneity between the sexes and between age groups. Age-related differences in the force of infection might arise from variation in the contribution of different routes of transmission, for example, an increased role for sexual transmission in young adults. Higher incidence in young males (age 0–4 years) may be indicative of risk-behaviour differences between the sexes restricted to early years of life. Such observations suggest areas to target research. The mechanism by which horizontal transmission in childhood is effected remains a phenomenon. We might speculate on cultural practices such as male circumcision as playing a role. Anthropological investigation of the gender differences in behaviour of young children is required to shed some light on the situation. Core groups who, through behavioural characteristics, are more likely to be infected and thus to infect others are thought to be of considerable importance to the persistence of HBV in low endemicity settings [13]. The results of this study show that also in the medium–high endemicity setting adult sexual transmission may be of importance, and further investigation of HBV seroprevalence in relation to patterns of sexual behaviour would be merited. It remains the case that higher incidence in young adults may arise through an increase in horizontal non-sexual transmission. There is a general lack of understanding of the routes of HBV transmission in adults in developing countries, highlighted by our observations. Finally, disentangling the role of social-behavioural characteristics from innate age or genetic factors, in generating patterns of age-prevalence is complicated. It is possible, for example, that the observed differential in carrier prevalence (HBsAg) and infectivity (HBeAg positive carriers), between the sexes (which may have a genetic or behavioural origin), could generate higher male seroprevalence, assuming, that is, there was significant within-gender contact rates. The reverse might also be the case, i.e. higher incidence of HBV in young males (where the risk of becoming a carrier is far higher than at older ages [48]), due to some sex difference in behaviour, could play a role in the higher male carrier prevalence.

The intrinsic capacity for spread of HBV in Addis Ababa, as defined by the basic reproductive rate, R_0, is estimated to be around 3–4. This represents a low potential relative to typical childhood contagious diseases such as measles and rubella [10, 30]. The difference in transmissibility of rubella and HBV, in the same population at the same time, is reflected also in the average age at infection (HBV >15 years; rubella 5 years) and A_{50} (HBV ~ 20 years; rubella 4 years) [30]. Our estimates of R_0 are based on simplifying assumptions about demography and mixing patterns, and are therefore only approximate. Demographically, Addis Ababa is a population undergoing marked decline in fertility, leading to progressive reduction in the proportional representation of children in the population. The impact on HBV endemicity of such demographic transition is thought to be complex-unidentifiable in the short term but potentially catastrophic in the long term [49], – and further study is warranted in this area given the general trend for demographic transition in the developing world. Furthermore, this recent theoretical work [49] suggests that similar levels of marker prevalence may arise from widely different values of R_0. However, the level of HBV prevalence identified from our study would remain consistent with values of $R_0 < 3$. In addition, an analysis based on a fully age-structured mathematical model of HBV transmission in The Gambia [15] gave an estimate of $R_0 \sim 2$ for this high endemicity country.

Thus calculations of the critical vaccination proportion for elimination of HBV in Addis Ababa of between 63–75% are unlikely to be underestimates. Given that perinatal transmission in this population is uncommon, it is likely that delivery of vaccine with routine DPT where current uptake is over 70% for three doses, would, in the long run be a successful control policy. A recent survey in Addis Ababa showed that of the increase in cost (40%) associated with an introduction of HBV vaccine into the EPI schedule, most of it (79%) would be the cost of vaccine (assumed to be US$0.5 per dose) [50]. If HBV vaccine costs continue to decline with growth of the world market, vaccine intervention will become increasingly cost-effective. However, Addis Ababa is not representative of Ethiopia in general, not so much in terms of the regional differences in HBV seroprevalence in Ethiopia, which are not dramatic [9, 20, 24], but in the level of routine vaccine coverage achieved, the national average being 51% in 2000/2001 with wide regional variation (personal communication, EPI, Ministry of Health, Addis Ababa). This does not necessarily argue against introducing HBV vaccine into the national EPI schedule, since low level coverage is predicted to have significant impact on carriage

prevalence [11], but rather that a national elimination policy would be unrealistic at present.

ACKNOWLEDGEMENTS

The study was made possible by the generous participation of Addis Ababa inhabitants and the support of the Ministry of Health, Region 14 Health Bureau, and Kebele Officials. In particular we thank Dr Eyob Tsegaye (previously head of Region 14 Health Bureau, Addis Ababa) and Dr Wondemagegnehu Alemu (previously EPI manager, Ministry of Health, Addis Ababa) for their support in conducting the survey work. The study had financial support from the Wellcome trust (Project grant no. 039056).

REFERENCES

1. Barin F, Perrin J, Chotard J, et al. Cross-sectional and longitudinal epidemiology of hepatitis B in Senegal. Progress Med Virol 1981; **27**: 148–162.
2. Hyams K, Okoth F, Tukei P, et al. Epidemiology of hepatitis B in Eastern Kenya. J Med Virol 1989; **28**: 106–109.
3. Tabor E, Bayley AC, Cairns J, Pelleu L, Gerety RJ. Horizontal transmission of hepatitis B virus among children and adults in five rural villages in Zambia. J Med Virol 1985; **15**: 113–120.
4. Miller W, Shao J, Weaver D, Shimokura G, Paul D, Lallinger G. Seroprevalence of viral hepatitis in Tanzanian adults. Trop Med Int Health 1998; **3**: 757–763.
5. Feret E, Larouze B, Diop B, Sow M, London WT, Blumberg BS. Epidemiology of heptatitis B virus infection in the rural community of Tip, Senegal. Am J Epidemiol 1987; **125**: 140–149.
6. Martinson F, Weigle K, Mushahwar I, Weber D, Royce R, Lemon S. Seroepidemiological survey of hepatitis B and C virus infections in Ghanaian children. J Med Virol 1996; **48**: 278–283.
7. Okoth FA, Kobayashi M, Kaptich DC, et al. Seroepidemiological study for HBV markers and anti-delta in Kenya. East Afr Med J 1991; **68**: 515–523.
8. Kiire CF, African RSG. Hepatitis B infection in sub-Saharan Africa. Vaccine 1990; **8**: s107–s112.
9. Tsega E, Mengesha B, Hansson BG, Lindberg J, Nordenfelt E. Hepatitis A, B, and delta infection in Ethiopia: a serological survey with demographic data. Am J Epidemiol 1986; **123**: 344–350.
10. Anderson RM, May RM. Infectious diseases of humans: dynamics and control. Oxford: Oxford University Press, 1991.
11. Edmunds WJ, Medley GF, Nokes DJ. The transmission dynamics and control of hepatitis B virus in The Gambia. Stat Med 1996; **15**: 2215–2233.
12. Edmunds WJ, Medley GF, Nokes DJ. Vaccination against hepatitis B virus in highly endemic areas: waning vaccine-induced immunity and the need for booster doses. Trans R Soc Trop Med Hyg 1996; **90**: 436–440.
13. Williams JR, Nokes DJ, Medley GF, Anderson RM. The transmission dynamics of hepatitis B in the UK: a mathematical model for evaluating costs and effectiveness of immunization programmes. Epidemiol Infect 1996; **116**: 71–89.
14. Williams JR, Nokes DJ, Anderson RM. Targeted hepatitis B vaccination – a cost effective immunisation strategy for the UK? J Epidemiol Commun Health 1996; **50**: 667–673.
15. Wilson JN, Nokes DJ, Carman WF. Predictions of the emergence of vaccine-resistant hepatitis B in The Gambia using a mathematical model. Epidemiol Infect 2000; **124**: 295–307.
16. Edmunds WJ, Medley GF, Nokes DJ, O'Callaghan CJ, Whittle HC, Hall AJ. Epidemiological patterns of hepatitis B virus (HBV) in highly endemic areas. Epidemiol Infect 1996; **117**: 313–325.
17. Tsega E, Nordenfelt E, Hansson BG, Mengesha B, Lindberg J. Chronic liver disease in Ethiopia: a clinical study with emphasis on identifying common causes. Ethiop Med J 1992; **30**: 1–33.
18. Tsega E. Epidemiology, prevention and treatment of viral hepatitis with emphasis on new developments. Ethiop Med J 2000; **38**: 131–141.
19. Weiland O. The impact of viral hepatitis on the morbidity and mortality of chronic liver disease and hepatocellular carcinoma in Ethiopia. Ethiop Med J 1992; **30**: i–iv.
20. Kefene H, Rapicetta M, Rossi GB, et al. Ethiopian National Hepatitis B Study. J Med Virol 1988; **24**: 75–84.
21. Nokes DJ, Enquselassie F, Nigatu W, et al. Has oral fluid the potential to replace serum for the evaluation of population immunity levels? A study of measles, rubella and hepatitis B in rural Ethiopia. Bull WHO 2001; **79**: 588–595.
22. Gebreselassie L. Occurrence of hepatitis B surface antigen and its antibody in Ethiopian blood-donors. Ethiop Med J 1983; **21**: 205–208.
23. Rapicetta M, Hailu K, Morace G, et al. Prevalence of HBeAg, anti-HBe serological markers and HBV-DNA in asymptomatic carriers in Ethiopia. Euro J Epidemiol 1989; **5**: 481–484.
24. Rahlenbeck SI, Yohannes G, Molla K, Reifen R, Assefa A. Infection with HIV, syphilis and hepatitis B in Ethiopia: a survey in blood donors. Int J STD AIDS 1997; **8**: 261–264.
25. Tsega E. Viral hepatitis during pregnancy in Ethiopia. East Afr Med J 1976; **53**: 270–277.
26. Tsega E, Hansson B-G, Krawczynski K, Nordenfelt E. Acute sporadic viral hepatitis in Ethiopia: causes, risk factors, and effects on pregnancy. Clinical Infect Dis 1992; **14**: 961–965.
27. Tsega E, Mengesha B, Nordenfelt E, Hansson BG, Lindberg J. Prevalence of hepatitis B virus markers

among Ethiopian blood donors: is HBsAg screening necessary? Trop Geogr Med 1987; **39**: 336–340.
28. Tsega E, Tsega M, Mengesha B, Nordenfelt E, Hansson BG, Lindberg J. Transmission of hepatitis B virus infection in Ethiopia with emphasis on the importance of vertical transmission. Int J Epidemiol 1988; **17**: 874–879.
29. Fontanet AL, Messele T, Dejene A, et al. Age- and sex-specific HIV-1 prevalence in the urban community setting of Addis Ababa, Ethiopia. Aids 1998; **12**: 315–322.
30. Cutts FT, Abebe A, Messele T, et al. Sero-epidemiology of rubella in the urban population of Addis Ababa, Ethiopia. Epidemiol Infect 2000; **124**: 467–479.
31. Population Housing and Census Commission. The 1994 population and housing census of Ethiopia. Results for Addis Ababa Vol. 1 – statistical report. Addis Ababa, Ethiopia: Central Statistical Authority, 1995.
32. Population Housing and Census Commission. The 1984 population census of Ethiopia. Results for Addis Ababa. Addis Ababa: Central Statistical Authority, 1985.
33. Bennett S, Woods T, Liyange WM, Smith DL. A simplified general method for cluster-sample surveys of health in developing countries. World Health Stat Q 1991; **44**: 98–106.
34. Smith P, Morrow R. Field trials of health interventions in developing countries: a toolbox. London: Macmillan Education Ltd, 1996: 362.
35. Nokes DJ, Forsgren M, Gille E, Ljungstrom I. Modelling longitudinal toxoplasma seroprevalence in Stockholm, Sweden. Parasitol 1993; **107**: 33–40.
36. McLean AR, Anderson RM. Measles in developing countries, Part I Epidemiological parameters and patterns. Epidemiol Infect 1988; **100**: 111–133.
37. McLean AR. Mathematical modelling of the immunisation of populations. Rev Med Virol 1992; **2**: 141–152.
38. Enquselassie F, Ayele W, Dejeue A, et al. Sero-epidemiology of measles in Addis Ababa, Ethiopia: implications for control through vaccination. Epidemiol Infect. In press.
39. Bowry TR. Seroepidemiology of Hepatitis B in an urban population of Nairobi, Kenya. J Infect Dis 1983; **148**: 1122.
40. Triki H, Said N, Ben Salah A, et al. Seroepidemiology of hepatitis B, C and delta viruses in Tunisia. Trans R Soc Trop Med Hyg 1997; **91**: 11–14.
41. Andre F. Hepatitis B epidemiology in Asia, the Middle East and Africa. Vaccine 2000; **18** (Suppl 1): S20–S22.
42. Maynard JE, Kane MA, Hadler SC. Global control of hepatitis B through vaccination: role of hepatitis B vaccine in the expanded programme on immunization. Rev Infect Dis 1989; **2**: S574–S579.
43. Colombo M, Gerber MA, Vernace SJ, Gianotti F, Paronetto F. Immune response to hepatitis B virus in children with papular acrodermatitis. Gastroenterol 1977; **73**: 1103–1106.
44. Coursaget P, Yvonnet B, Chotard J, et al. Age- and sex-related study of hepatitis B virus chronic carrier state in infants from an endemic area (Senegal). J Med Virol 1987; **22**: 1–5.
45. McMahon BJ, Alward WLM, Hall DB, et al. Acute hepatitis B virus infection: relation of age to the clinical expression of disease and subsequent development of the carrier state. J Infect Dis 1985; **151**: 599–603.
46. Kashiwagi S, Hayashi J, Nomura H, Ikematsu H, Kajiyama W. Large-scale survey of hepatitis B virus infection in families. Microbiol Immunol 1985; **29**: 951–958.
47. Sobeslavsky O. Prevalence of markers of hepatitis B virus infection in various countries: a WHO collaborative study. Bull WHO 1980; **58**: 621–628.
48. Edmunds WJ, Medley GF, Nokes DJ, Hall AJ, Whittle HC. The influence of age on the development of the hepatitis B carrier state. Proc R Soc Lond B Biol Sci 1993; **253**: 197–201.
49. Medley GF, Lindop NA, Edmunds WJ, Nokes DJ. Hepatitis-B virus endemicity: heterogeneity, catastrophic dynamics and control. Nat Med 2001; **7**: 619–624.
50. Edmunds W, Dejene A, Mekonnen Y, Haile M, Alemnu W, Nokes D. The cost of integrating hepatitis B virus vaccine into national immunization programmes: a case study from Addis Ababa. Health Policy Plan 2000; **15**: 408–416.