

# THE LANCET

## **Supplementary webappendix**

This webappendix formed part of the original submission and has been peer reviewed. We post it as supplied by the authors.

Supplement to: Scott JAG, Berkley JA, Mwangi I, et al. Relation between falciparum malaria and bacteraemia in Kenyan children: a population-based, case-control study and a longitudinal study. *Lancet* 2011; published online Sept 7. DOI:10.1016/S0140-6736(11)60888-X.

# **Webappendix for: Falciparum malaria is a major risk factor for bacteraemia in Kenyan children; evidence from the malaria-protective effect of sickle cell trait.**

## **Methods of the case-control studies**

### **1. Matched case-control study**

The objective of this study was to describe the major biological predisposing factors for the invasive bacterial disease in Kenyan children, including HIV, sickle cell disease, malaria and malnutrition. We selected hospitalisation with bacteraemia as a case-definition for invasive bacterial disease because it was specific and widely recognised, and because we were planning simultaneously to do an inpatient study of the aetiology of bacteraemia among children in Kilifi (1). Both studies were reviewed and approved by the KEMRI National Ethical Review Committee within the same protocol; each study had separate but overlapping study populations and had separate consent forms.

#### **Case definition**

Children were eligible for recruitment in the case-control study if they were admitted to Kilifi District Hospital between 16th September 1999 and 31st of July 2002. During this period, under the provisions of the bacteraemia aetiology study, all children admitted to the paediatric ward were investigated with blood cultures on admission except those with elective surgery or minor trauma requiring overnight observation. On admission, each child's skin was cleaned with 70 percent ethanol and allowed to dry before blood was drawn for culture and antimicrobial testing. Blood was inoculated into a culture bottle (BACTEC Peds Plus, Becton Dickinson) after the needle had been changed and the top had been cleaned with 70 percent ethanol and allowed to dry. If the blood cultures taken on admission yielded a pathogenic bacterial isolate, the microbiology laboratory informed the clinical research staff who adjusted treatment where necessary and approached the family of the patient to ask for consent for the participation of their child in the case-control study. All bacteria cultured from the blood were considered pathogenic except those commonly found as contaminants, including coagulase-negative staphylococci, micrococcus species, bacillus species and viridans streptococci.

#### Exclusion criteria:

- a. Age <3 months. Young infants have a different pattern of pathogens causing invasive bacterial disease and are likely to have a different set of risk factors.
- b. Age >13 years. This was the upper limit for admission to the paediatric ward.
- c. Discharged from hospital within the last 14 days. The objective of the study was to define the risks of community-acquired invasive bacterial disease not hospital-acquired infection.
- d. Death in hospital between admission and the isolation of a pathogenic bacterium. Ascertainment of the full list of exposures would be incomplete if the child could not be examined by the research team.
- e. Absconding from the hospital between admission and the isolation and identification of a pathogenic bacterium.

#### **Control definition**

Healthy community controls were individually matched to cases on age, sex, geographic location and time. Wherever possible, two controls were recruited for each case identified. At the point a case was recruited, a fieldworker was notified of the requirement to recruit two control participants. The fieldworker obtained the directions to the child's home by interviewing the relatives caring for the child in hospital and then travelled to the house. There, he spun a pencil on the ground to provide an random direction in which to seek out a location-matched control. Continuing always in this direction, he asked that each house in turn whether the house contained a resident child who fitted the matching criteria of sex and age; the control child should be the same sex as the case and should fall in the same age stratum. The predefined age strata were 2-11, 12-23, 24-59 and  $\geq 60$  months. Having found a matched control, the fieldworker continued walking in the same direction until he was able to recruit a second control. The fieldwork was required to recruit two controls within 14 days of the identification of the case because several of the exposure variables were strongly associated with season.

### Exclusion criteria:

- a. Aged <2 months.
- b. Aged >13 years.
- c. Discharge from hospital within the last 14 days.
- d. Child sick on the day of recruitment.

### **Exposure ascertainment**

At the first contact between the research staff and a case, in hospital, or a control at his/her own home, the study was explained to the parents/guardians of the child and written consent was obtained for participation in the study. Exposure variables obtained by questionnaire were ascertained by interviewing the parents/guardians of cases on the ward and of controls in their own homes. Risk factors that required clinical examination were ascertained at the hospital; the parents of healthy controls were given an appointment to attend the hospital and the child was examined by the study doctor and underwent a blood test. The questionnaire variables included age, sex, current health status, location of residence, date of recruitment, number of years of education of father and mother, mother's ethnic group and father's ethnic group. At clinical examination the child was weighed and measured (height or length as appropriate) using scales and rulers that were calibrated weekly. A mid upper arm circumference was taken using a tape measure designed for the purpose. The study doctor examined the child for skin and hair changes of malnutrition, pedal oedema, a BCG scar and splenomegaly. Blood was taken for haemoglobin concentration, red and white blood cell counts, a malaria blood film, HRP2 assay, the presence of leucocyte haemozoin pigment and determination of sickle cell status. Blood cultures were not obtained from controls.

### **Laboratory methods**

Blood was cultured using a BACTEC 9050 automated instrument (Becton Dickinson, UK). Bottles that signalled positive were sub-cultured onto blood agar, chocolate agar and McConkey agar. Bottles were considered negative if they did not signal within 5 days, and were discarded. Isolates growing on these sub-cultures were identified using conventional microbiological techniques and biochemical tests (API, Biomerieux). Quality assurance was provided by the UK National External Quality Assessment Service ([www.ukneqas.org.uk](http://www.ukneqas.org.uk)). Isolates of *Streptococcus pneumoniae* were serotyped by Quelling reaction and of *Haemophilus influenzae* were serotyped by latex agglutination.

Antibodies to HIV were tested by enzyme-linked immunosorbent assay (ELISA) (Vironostika, Biomerieux) and dipstick (Determine, Abbott Laboratories). Positive samples from children under 18 months of age and samples with discordant results were assayed by PCR for proviral DNA. Haemoglobin, red blood cell count and white blood cell count were assayed by Coulter counter. Blood was examined for malaria parasites by microscopy of Giemsa-stained thick and thin films. Parasite densities were calculated as parasite counts per 500 red blood cells (or per 100 white blood cells) multiplied by the relevant blood cell concentration measured by Coulter counter. Leucocyte haemozoin pigment, a marker of recent malaria infection, was detected by slide microscopy (2). Plasma levels of the malaria antigen histidine rich protein 2 (HRP2) were measured by sandwich ELISA against a known concentration standard (3); HRP2 is secreted by red blood cells infected by *P. falciparum* and is detectable in the plasma of adult patients with uncomplicated malaria for  $\geq 2$  weeks (4). HbS phenotype was determined by citrate-agar electrophoresis using standard methods.

### **Sample size**

The sample size was based on the estimated prevalences of the four major exposure variables. With alpha of 0.05, an assumed correlation between case-control pairs of 0.2 and a control/case ratio of 2, a study of 300 cases provided adequate power to detect a useful range of odds ratios across the prevalences of these variables.

Risk Factor	Estimated Prevalence	Odds Ratio			
		1.5	2.0	2.5	3.0
HIV	0.03	0.178	0.453	0.724	0.895
Malnutrition	0.07	0.321	0.764	0.960	0.996
Malaria pigment	0.1	0.411	0.873	0.990	1.000
Parasitaemia	0.3	0.709	0.992	1.000	1.000

The table shows the power of the study for exposure variables of different prevalence at a range of potential odds ratios varying between 1.5-3.0. Power calculations were derived from equations in Dupont (5)

### **Statistical analysis**

Data were assembled in case-control sets consisting of either pairs (where only one control was recruited) or triplets (where two controls per case were recruited). For each exposure variable the number of concordant and discordant pairs was calculated and tabulated. Odds ratios, 95% confidence limits and P values were calculated using classical methods (6). For the multivariable analysis, we performed hierarchical modelling using logistic regression, conditional on the case-control sets. For two key exposures, malaria and malnutrition, we had a variety of different measures of the exposure, and we wished to select the measure which best represented the exposure variable before controlling for confounding by other variables. We first repeated the univariate analysis using conditional logistic regression. Next, we developed separate intermediate models of these two sets of variables using backward stepwise regression. Each variable was tested with a likelihood ratio test and variables that were significant at  $P < 0.1$  were retained in the intermediate model. At the second stage of modelling, we included all variables with  $P < 0.1$  in the univariate analysis plus those retained in the intermediate models. We developed the final model using backward stepwise conditional logistic regression. Variables were retained in the final model if they had a likelihood ratio test  $P < 0.05$ . Finally, we examined the final model variables in two subsets of the cases-control sets, those which included a case of pneumococcal bacteraemia and those which included a case of non-Typhoidal salmonella.

### **2. Longitudinal case-control study**

This study was conceived and designed at the end of the study period in order to examine, specifically, the effect of declining malaria incidence on the protection against bacteraemia afforded to children with sickle cell trait. The opportunity to conduct a retrospective, longitudinal, case-control study arose out of two existing longitudinal studies: (1) a longitudinal clinical surveillance study of severe diseases of children, including bacteraemia, at Kilifi District Hospital and (2) a genetic cohort study of susceptibility to malaria conducted in the area of the Kilifi Health and Demographic Surveillance System (KHDSS), which surrounds Kilifi District Hospital (7).

### **Case definition**

We defined a case as any child, aged  $< 14$  years, admitted to Kilifi District Hospital between 1<sup>st</sup> January 1999 and 31<sup>st</sup> of December 2007 in whom a blood culture yielded a pathogenic bacterial isolate. Throughout this period children were investigated in a consistent manner with a blood culture on admission, as described in the matched case-control study. A child could only become a case if the pathogenic isolate was cultured from a blood sample that had been taken on admission; hospital-acquired infections were not included. The venesection method was constant throughout the period and is described under the section on the matched case-control study. All isolates of bacteria cultured from blood were considered pathogenic unless they were of species defined to be contaminants (listed above). As there was an overlap between the study periods for the matched case-control study and the longitudinal case-control study, some cases were included in both studies. However, in the longitudinal case-control study the number of exposures ascertained was much smaller than in the matched study and these were readily available within the clinical surveillance study database. Therefore it was not necessary to exclude children because they died between admission to hospital and the isolation of a pathogen from blood.

### **Exclusion criteria**

- A. Age  $> 13$  years
- B. Not resident within the KHDSS area
- C. Did not consent to the Kilifi District Hospital clinical surveillance study.
- D. Already recruited to the study within the last 6 months.

### **Control definition**

For this study we used as controls a cohort of children who had been recruited into a study of genetic susceptibility to malaria and other infectious diseases. The Kilifi Genetic Birth Cohort Study (KGBCS) began recruiting children on 1 January 2006 and we included in this case-control analysis all children recruited up until 23 June 2009. The objective of the cohort recruitment procedure was to approach

every child aged 3 to 11 months who was resident in the Kilifi Health and Demographic Surveillance Study area. The fieldworkers recruiting participants in this study obtained lists of residents, including newly born children, from the KHDSS and visited the family at home to explain the study and to invite the parents/guardians to consent to the participation of their child. Because of the nature of the study, looking at genetic susceptibility traits, the research team developed a comprehensive plan of community communication before embarking on the study. The results of the major genotype under scrutiny, sickle cell disease, were communicated back to the parents/guardians.

#### Exposure ascertainment

In this study only four exposure variables were analysed; sex, ethnic group, geographic location of residence, and sickle cell status. Among cases, sex, ethnic group and geographic location were recorded in the computer-based clinical records of the Kilifi clinical surveillance study. We examined sickle-cell status by assaying stored whole blood samples that had been frozen at the point of admission to hospital. Among controls, sex, ethnic group and geographical location were already known at the point of recruitment, from the listing of residents in the KHDSS. Sickle cell status was ascertained from capillary blood samples obtained at home visits shortly after recruitment to the cohort study, and analysed for sickle cell status within 7 days of collection.

#### **Laboratory methods**

Blood cultures were processed and analysed in a consistent manner throughout the whole of the study period; the methods are described in detail under the matched case-control study.

DNA from cases was extracted, retrospectively, from frozen blood samples, collected on admission, using Qiagen DNA blood mini kits (Qiagen, Crawley, UK) and typed for HbS by PCR (8). For controls, capillary blood samples were phenotyped for HbS by HPLC (Variant Analyzer, BioRad, Hercules, CA) using the  $\beta$ -thalassaemia short program.

#### **Statistical analysis**

The sample size in this study was determined by the number of cases and controls available within the existing Kilifi clinical surveillance records and the KGBCS rather than by theoretical consideration or calculation. We used logistic regression to estimate the odds ratio and 95% confidence limits for sickle cell trait among cases and controls, taking account of potential confounders sex, ethnic group, and geographical residence within the Kilifi Health and Demographic Surveillance Study area. To control for geographic residence we divided the area into five strata. In order to observe the change in the odds ratio over time we conducted a separate analysis for each year of the nine-year study period; all of the cases identified in one year were compared against all of the controls within the KGBCS. This approach produced nine odds ratios describing the trend in the association between sickle cell trait and bacteraemia during the period of declining malaria incidence.

Two limitations of this method are noteworthy. Firstly, we were unable to control for age in this analysis; although the cases varied in age from newborns to 13-year-olds, all of the controls were infants. Whilst we know that the prevalence of sickle cell disease declines sharply in early childhood, the prevalence of sickle cell trait does not increase appreciably through early childhood. It is very unlikely therefore that this led to a perceptible degree of confounding. Secondly, we were unable to control for time; although the cases were recruited evenly throughout a nine-year period, the controls were all recruited within a three-year period at the end of the study. It is possible, that the decline in the burden of malaria in this area during the study period could have led to a decline in the gene frequency of the  $\beta^s$  gene and this could have contributed to the change in the odds ratio of sickle cell trait and bacteraemia. The potential for such confounding can be seen and evaluated in the prevalence of sickle cell trait among the controls in the matched case-control study and the longitudinal case-control study. A small change is unlikely to lead to confounding. Furthermore, a comparison of the odds ratios obtained at the end of the longitudinal case control study with those obtained in the matched case-control study will illustrate the potential for confounding by time, because in each instance the controls were recruited contemporaneously with the cases.

## Supplementary Tables and Figures

**Table S1. Pathogenic bacteria isolated from blood cultures among children admitted to Kilifi District Hospital between 1999 and 2002 who were recruited as cases in the matched case-control study.**

<b>Organism</b>	<b>Frequency</b>	<b>Percent</b>
<i>Acinetobacter</i> sp	18	6.16
<i>Pantoea agglomerans</i>	2	0.68
<i>Escherichia coli</i>	19	6.51
<i>Enterobacter</i> sp	2	0.68
<i>Haemophilus influenzae</i>	37	12.7
<i>Klebsiella pneumoniae</i>	3	1.03
<i>Neisseria meningitidis</i>	1	0.34
<i>Pseudomonas aeruginosa</i>	2	0.68
<i>Pseudomonas oryzae</i>	1	0.34
<i>Pseudomonas stutzeri</i>	2	0.68
<i>Salmonella</i> sp (non-Typhi)	61	20.9
<i>Staphylococcus aureus</i>	15	5.14
<i>Shigella flexneri</i>	2	0.68
haemolytic <i>Streptococcus</i> Group A	5	1.71
haemolytic <i>Streptococcus</i> Group B	1	0.34
<i>Enterococcus</i> spp	7	2.40
haemolytic <i>Streptococcus</i> Group G	2	0.68
<i>Streptococcus pneumoniae</i>	111	38.0
<i>Salmonella</i> Typhi	1	0.34
<b>Total</b>	<b>292</b>	<b>100</b>

**Table S2. Numbers of cases and control pairs, concordant and discordant, in the classical analysis of the matched case-control study together with classical univariate odds ratios.**

exposure variable	case	number of exposed controls					No. of pairs	No. of discordant pairs	Analytic results				
		CC pairs		CC triplets					OR	lower CI	upper CI	p	
	status	0	1	0	1	2							
HbSS (v HbAA)	unexp	87	1	138	0	0	383	20	15	2.3	97.91	0.0007	
	exp	7	0	6	0	0							
HbAS (v HbAA)	unexp	35	9	138	52	7	483	111	0.47	0.29	0.77	0.003	
	exp	3	2	13	7	0							
HIV infection	unexp	39	2	174	6	0	527	132	14.56	7.10	29.84	<0.0005	
	exp	14	0	54	2	0							
<i>P. falciparum</i> >0	unexp	33	14	78	62	33	528	219	0.53	0.38	0.76	<0.0005	
	exp	5	4	25	22	16							
<i>P. falciparum</i> >10,000/mcl	unexp	51	4	181	34	3	528	78	0.75	0.43	1.31	0.48	
	exp	1	0	15	3	0							
<i>P. falciparum</i> >50,000/mcl	unexp	54	1	215	8	0	528	36	2.89	1.25	6.70	0.03	
	exp	1	0	13	0	0							
Haemozoin pigment	unexp	46	0	169	23	2	524	111	3.28	2.02	5.31	<0.0005	
	exp	9	1	35	5	0							
HRP2>0ng/ml	unexp	79	6	57	4	1	254	52	3.03	1.52	6.07	0.002	
	exp	11	2	13	3	0							
HRP2>20ng/ml	unexp	90	1	74	1	0	251	12	4.8	0.99	23.24	0.14	
	exp	4	0	3	0	0							
palpable spleen	unexp	33	4	113	50	6	528	194	2.01	1.43	2.81	<0.0005	
	exp	17	2	47	17	3							
weight-for-age z-score ≥-1.8	unexp	16	21	24	75	44	528	257	0.39	0.28	0.54	<0.0005	
	exp	5	14	15	38	40							
height-for-age z-score ≥-1.6	unexp	13	15	38	59	33	528	241	0.71	0.52	0.96	0.043	
	exp	7	21	24	46	36							
weight-for-height z-score ≥-1.1	unexp	15	25	36	66	42	528	252	0.43	0.31	0.59	<0.0005	
	exp	6	10	12	47	33							
skin or hair changes of malnutrition	unexp	44	2	180	12	1	528	106	5.56	3.15	9.81	<0.0005	
	exp	9	1	38	5	0							
pedal oedema	unexp	55	0	212	0	0	530	51	-	-	-		
	exp	3	0	24	0	0							
BCG scar	unexp	1	4	1	10	34	525	155	0.94	0.65	1.36	1	
	exp	12	40	6	49	134							
maternal ethnic group not Mijikenda	unexp	46	3	210	10	1	524	45	2.03	1.01	4.08	0.12	
	exp	7	0	11	1	1							
paternal ethnic group not Mijikenda	unexp	47	3	212	11	1	524	36	1.29	0.61	2.7	1	
	exp	5	1	6	3	1							
maternal education ≥1 year	unexp	15	13	46	52	25	514	228	0.99	0.73	1.34	1	
	exp	14	12	25	49	33							
paternal education ≥7 years	unexp	15	18	11	31	24	382	166	0.7	0.49	1	0.09	
	exp	11	36	9	40	36							

**Table S3. Subgroup analysis of the matched case-control study to examine specific risk factors for pneumococcal bacteraemia and non-Typhi Salmonella bacteraemia.**

Exposure Variable	All bacteraemia cases		Pneumococcal bacteraemia		non-Typhi Salmonella bacteraemia	
	OR	95%CI	OR	95%CI	OR	95%CI
HbSS	22.9	2.47, 211	-	-	9.35	0.57, 153
HbAS	0.36	0.20, 0.65	0.35	0.13, 0.91	0.21	0.05, 0.97
HIV	9.03	3.59, 22.7	14.9	3.28, 67.7	11.3	1.32, 96.4
Leucocyte haemozoin pigment	3.52	1.92, 6.47	1.51	0.59, 3.86	16.5	3.44, 79.3
Mid-upper arm circumference/cm	0.57	0.49, 0.66	0.67	0.53, 0.84	0.43	0.27, 0.68

Odds ratios and 95% confidence intervals were estimated using conditional logistic regression with the subset of cases with *S. pneumoniae* bacteraemia and non-Typhi Salmonella bacteraemia and their matched controls. The *S. pneumoniae* analysis contained 104 cases and 202 controls; the Salmonella analysis contained 60 and 109 controls. Exposure variables selected were those that were significant in the final adjusted model of the full dataset (see Table 1). The odds ratio for HbSS is missing in the pneumococcal sub-group model because none of the cases of pneumococcal bacteraemia had sickle cell disease (HbSS).

**Table S4. Distributions of sickle cell trait and leucocyte haemozoin pigment among cases who were bacteraemic with different pathogen subgroups.**

Exposure	Bacterial pathogen isolated in blood cultures				Fisher Exact P value
	<i>Streptococcus pneumoniae</i>	non-Typhi Salmonella	Other Gram positive	Other Gram negative	
HbAA	92	52	18	84	
HbAS	9	4	4	8	0.49
no pigment	98	41	21	82	
leucocyte haemozoin pigment	13	20	1	15	0.003

The results of haemoglobin electrophoresis were unavailable from 7 bacteraemic cases and a further 14 cases were excluded because they had sickle cell disease (HbSS). The table tests the hypotheses that (i) the association between sickle cell trait and bacteraemia varies by bacterial pathogen and (ii) the association between leucocyte haemozoin pigment and bacteraemia varies by pathogen. This is found to be true for leucocyte haemozoin pigment but not for sickle cell phenotype.



**Table S5. Population size, numbers of paediatric admissions to Kilifi District Hospital with malaria or bacteraemia and incidence of malaria and bacteraemia 1999-2007.**

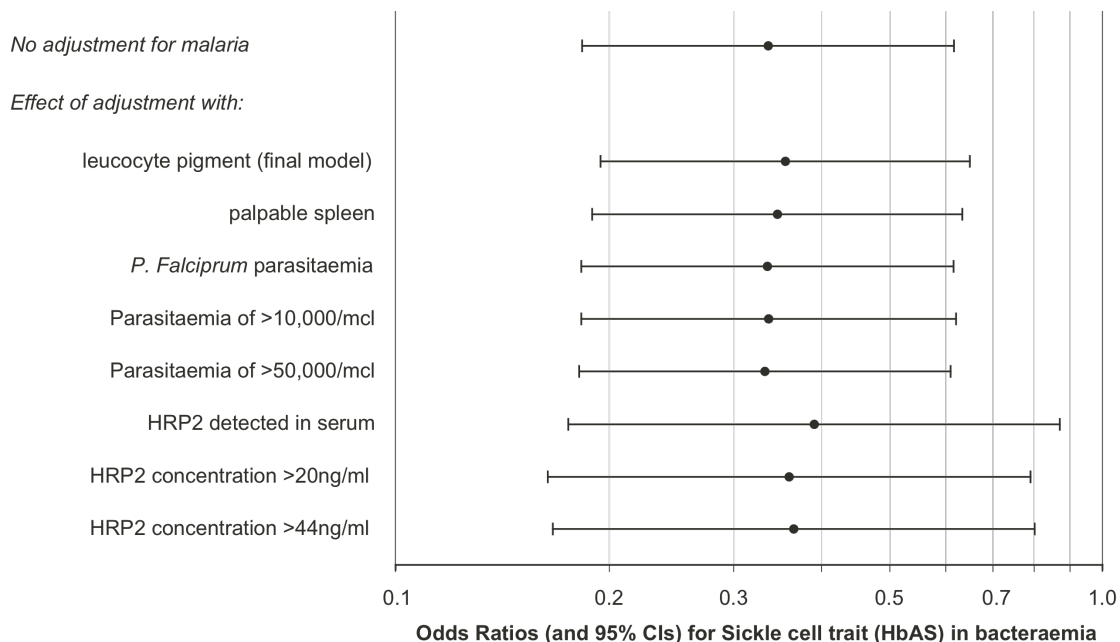
Year	1999	2000	2001	2002	2003	2004	2005	2006	2007
Population <14 years	76,210	81,033	86,161	91,614	97,991	102,280	106,046	108,948	111,054
<b>Malaria slide positive admissions</b>									
Total	2,170	1,898	1,802	1,617	1,799	1,145	770	623	383
Incidence*	28.47	23.42	20.91	17.65	18.36	11.19	7.26	5.72	3.45
Lower CI <sup>#</sup>	27.29	22.38	19.96	16.80	17.52	10.56	6.76	5.28	3.11
Upper CI <sup>#</sup>	29.70	24.50	21.90	18.53	19.23	11.86	7.79	6.19	3.81
<b>Bacteraemia positive admissions</b>									
Any pathogenic bacteria	197	200	197	181	231	200	161	158	161
Incidence*	2.59	2.47	2.29	1.98	2.36	1.96	1.52	1.45	1.45
Lower CI <sup>#</sup>	2.24	2.14	1.98	1.70	2.06	1.69	1.29	1.23	1.23
Upper CI <sup>#</sup>	2.97	2.83	2.63	2.29	2.68	2.25	1.77	1.69	1.69
<b>Gram positive organisms</b>									
<i>S pneumoniae</i>	40	64	63	57	84	73	51	39	51
<i>S aureus</i>	18	8	23	15	26	24	18	25	24
Group A Streptococci	11	7	14	13	14	12	7	7	5
Other Gram positive organisms <sup>†</sup>	9	12	19	13	13	12	25	21	13
Any Gram positive organisms	77	91	118	97	135	115	100	92	91
<b>Gram negative organisms</b>									
<i>H influenzae</i> non-type b	6	7	2	3	5	3	4	3	1
non-Typhi Salmonella	52	32	25	37	34	13	10	8	13
<i>E coli</i>	22	20	14	14	20	15	8	18	21
Acinetobacter species	16	17	17	14	22	26	15	20	18
Pseudomonas species	8	13	6	7	6	9	11	9	3
Klebsiella species	4	8	9	3	1	9	5	0	3
Other Gram negative organisms <sup>‡</sup>	17	16	14	10	11	15	14	11	11
Any Gram negative organism	123	110	85	88	97	89	67	69	70

Except where marked (e.g. \*Incidence / 1000 / year; <sup>#</sup>95% confidence intervals) numbers in cells are counts of children admitted to hospital with bacteraemia of the categories indicated. One child with *H. influenzae* type b bacteraemia also had pneumococcal bacteraemia (the child was excluded from this analysis). Fifty children had 2 bacterial isolates and two children had 3 isolates from the same blood culture. Two of the children with multiple cultures had 2 "Other Gram negative organisms"; 28 children had at least one Gram positive and one Gram negative isolate from the same blood culture.

<sup>†</sup> includes Enterococcus species, Streptococcus groups B, C, F and G and *Staphylococcus saprophyticus*.

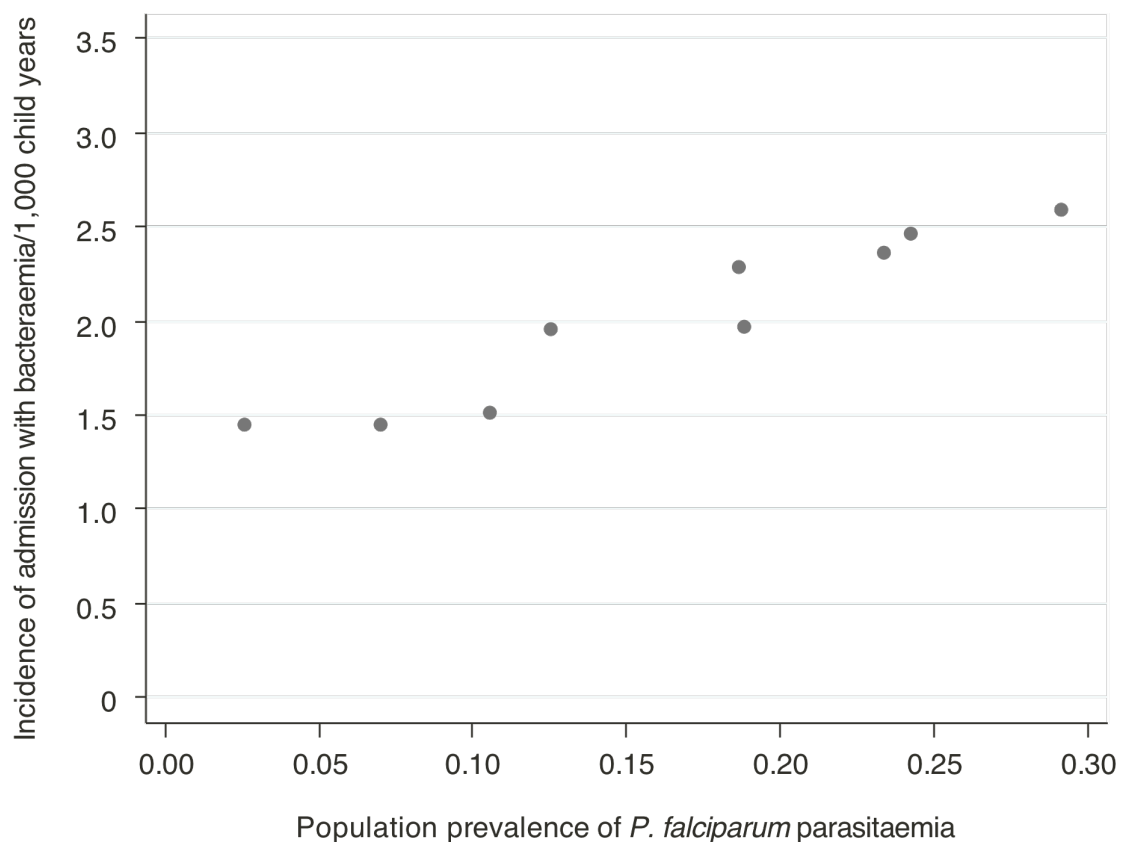
<sup>‡</sup> includes Campylobacter, Enterobacter, Neisseria, Proteus, *Salmonella* Typhi and Shigella species.

**Figure S1. Effect of individual adjustment for markers of malaria on the unadjusted odds ratio for the association between sickle cell trait (HbAS) and bacteraemia in children in the matched case-control study.**



Each line shows a different conditional logistic regression model of the matched case-control data after adjustment for HbSS, HIV and mid-upper arm circumference. Testing the hypothesis that malaria is on the causal pathway of the protective effect of HbAS on bacteraemia, the figure illustrates the effect of individual adjustment for malaria, using all the available markers in turn, on the odds ratio for HbAS. As accurate ascertainment of clinical malaria is difficult, especially among febrile IBD patients who may have taken anti-malarial drugs prior to presentation to hospital, we decided to control for malaria at the population level rather than the individual level in the longitudinal case-control study.

**Figure S2. Bacteraemia incidence each year among children aged <14 years admitted to Kilifi District Hospital from within the KHDSS study area against yearly prevalence of *P. falciparum* parasitaemia from 1999-2007.**



Episodes of *H. influenzae* type b bacteraemia are not included in the estimates of bacteraemia incidence. These data were used in a Poisson regression analysis to estimate the bacteraemia Incidence Rate Ratio for malaria parasitaemia.

## References

1. Berkley JA, Lowe BS, Mwangi I, Williams T, Bauni E, Mwarumba S, et al. Bacteremia among children admitted to a rural hospital in Kenya. *N Engl J Med*. 2005;352(1):39-47.
2. Nguyen PH, Day N, Pram TD, Ferguson DJ, White NJ. Intraleucocytic malaria pigment and prognosis in severe malaria. *Trans R Soc Trop Med Hyg*. 1995;89(2):200-4.
3. Ochola LB, Marsh K, Lowe B, Gal S, Pluschke G, Smith T. Estimation of the sequestered parasite load in severe malaria patients using both host and parasite markers. *Parasitology*. 2005;131(Pt 4):449-58.
4. Mayxay M, Pukrittayakamee S, Chotivanich K, Looareesuwan S, White NJ. Persistence of *Plasmodium falciparum* HRP-2 in successfully treated acute falciparum malaria. *Trans R Soc Trop Med Hyg*. 2001;95(2):179-82.
5. Dupont WD. Power calculations for matched case-control studies. *Biometrics*. 1988;44(4):1157-68.
6. Breslow NE, Day NE. *Statistical Methods in Cancer Research. Vol 1. The Analysis of Case-control Studies*. Lyon: International Agency for Research on Cancer; 1980.
7. Williams TN, Uyoga S, Macharia A, Ndila C, McAuley CF, Opi DH, et al. Bacteraemia in Kenyan children with sickle-cell anaemia: a retrospective cohort and case-control study. *Lancet*. 2009;374(9698):1364-70.
8. Waterfall CM, Cobb BD. Single tube genotyping of sickle cell anaemia using PCR-based SNP analysis. *Nucleic Acids Res*. 2001;29(23):E119.