

The epidemiology of periportal fibrosis and relevance of current *Schistosoma mansoni* infection: a population-based, cross-sectional study

Authors

Seun Anjorin¹PhD, Betty Nabatte² BA, Simon Mpooya² MSc, Benjamin Tinkitina² MSc, Christopher K. Opio³ MD, Narcis B. Kabatereine² PhD, Goylette F. Chami^{1*} PhD

Affiliations

1. Big Data Institute, Nuffield Department of Population Health, University of Oxford, Oxford, UK
2. Division of Vector-Borne and Neglected Tropical Diseases Control, Uganda Ministry of Health, Kampala, 15 Bombo Road, Kampala, Uganda
3. Aga Khan University Hospital, 3rd Parkland Avenue, PO Box 30270-00100, Nairobi, Kenya

*Correspondence

Big Data Institute
University of Oxford
Oxford, United Kingdom OX3 7LF
goylette.chami@ndph.ox.ac.uk
[+441865743598](tel:+441865743598)

Keywords

Fibrosis; hepatitis; HIV; periportal; Symmer's; pipestem; schistosomiasis; intestinal; hepatic; liver; *Schistosoma mansoni*; portal hypertension; Niamey; ultrasound; morbidity; Africa; SchistoTrack

ABSTRACT (277/300 words)

Background: Intestinal schistosome infections are known to cause periportal fibrosis (PPF). Yet, the epidemiology of PPF remains poorly understood, especially in settings endemic with *Schistosoma mansoni*.

Methods: We randomly sampled 1442 households from 38 villages in Mayuge, Buliisa, and Pakwach Districts of Uganda within the SchistoTrack Cohort to examine 2834 individuals aged 5-90 years. PPF was diagnosed using ultrasound and image patterns C-F from the Niamey Protocol. *S. mansoni* infection status/intensity was diagnosed by Kato-Katz microscopy and point-of-care circulating cathodic antigens (POC-CCA). Schistosome infection, coinfections, and comorbidities were examined as exposures for PPF. Logistic regressions were run with standard errors clustered by household.

Findings: PPF prevalence was 12·10% (343/2834), varying from 5·00-19·46% across districts. *S. mansoni* prevalence by Kato-Katz, and POC-CCA trace negative and positive was 43·37% (1229/2834), 40·86% (1158/2834), and 65·73% (1863/2834) respectively. Individual schistosome infection status/intensity was not correlated with the likelihood of PPF. Living in a village where adults had <5% prevalence of heavy intensity infections (400+ eggs per gram of stool) was associated with 30.2% decreased odds of PPF. The likelihood of PPF with age linearly increased from 5-25, exponentially changed from 26-45, remain unchanged from 45-60, and steadily decreased past 60 years. History of liver diseases, human immunodeficiency virus positivity (HIV+), and ultrasound-detected chronic hepatitis/early cirrhosis-like disease were associated with >2-fold increased PPF likelihood.

Interpretation: Current individual schistosome infections alone are uninformative for PPF. History of HIV+ and underlying chronic hepatitis/early cirrhosis-like disease were risk factors and could be investigated for PPF surveillance and management.

Funding: Nuffield Department of Population Health Pump Priming Fund, Wellcome Trust Institutional Strategic Support Fund (204826/Z/16/Z), John Fell Fund, Robertson Foundation Fellowship, and UKRI EPSRC Award (EP/X021793/1).

Research in context

Evidence before this study

Morbidity due to parasitic infection is a complex interplay of current and past exposures. World Health Organization (WHO) guidelines for elimination of schistosomiasis as a public health problem assume current infection is a reliable proxy indicator of prevalent morbidity. Community infection thresholds are defined in guidelines and, when met, the assumption is there is no schistosomiasis-related morbidity. There is a lack of evidence for the association of infection with prevalent morbidity in the context of repeated treatment from routine mass drug administration. To evaluate WHO guidelines, there is a need for large-scale population-based, cross-sectional studies in endemic areas where current infection is compared with current morbidity at the same timepoint. A cross-sectional design also enables the investigation of a wide range of risk factors to assess the relative importance of current schistosome infection. Periportal fibrosis is a schistosomiasis-associated severe morbidity with clinical consequences such as portal hypertension, upper gastrointestinal tract bleeding, and ultimately premature death. Yet, little is known about the distribution of this disease; no information is available on its most basic epidemiology including age and gender-specific likelihoods. It is schistosomiasis-specific or attributable to schistosome pathology unlike more subtle conditions with complex aetiologies (e.g. anaemia). Hence, investigating periportal fibrosis serves as a first line, conservative approach to evaluating World Health Organization guidelines for elimination of schistosomiasis-related morbidity as a public health problem. A systematic literature search as part of an ongoing metanalysis was prospectively registered on PROSPERO (CRD42022333919). Databases were searched on 18th May 2022 and included the Cochrane Central Register of Controlled Trials, Embase, Global Health, Global Index Medicus, and Medline. The following general terms were used: “Schistosoma” AND “fibrosis” AND “intensity” AND “infection” AND “periportal OR liver”. Studies were included that considered *Schistosoma mansoni*, *S. japonicum*, or *S. mekongi* species. Only original research articles in English were considered. No restriction on date of publication, age, gender, or region was applied. Infection was required to be diagnosed as opposed to self-reported. Periportal fibrosis was defined by the study authors.

Added value of this study

No population-wide or adjusted analyses were found to characterize the age-specific likelihood of periportal fibrosis. Studies focused on unadjusted associations in nonrandomly sampled populations of narrow age groups. In a single case (Weigand et al 2021) where adjusted analyses were completed (with age and gender considered), national programmatic data was used with sparsely nonrandomly sampled schoolchildren and limited frequency of the outcome of periportal fibrosis. Ultrasound data collection protocols and validation across studies were poorly reported. There was a lack of investigations on coinfections and comorbidities with only eight studies initiated (or published) from 2003 onwards after the start of mass drug administration in sub-Saharan Africa.

Implications of all available evidence

To our knowledge, this study is the first to characterize the epidemiology of periportal fibrosis with respect to the most common intestinal schistosome pathogen (*S. mansoni*). We conducted a comprehensive, population-based study of all ages (5+ years) eligible for mass drug administration in an area that has received at least 13 rounds of treatment. Here we provide clear evidence for the lack of association of current *S. mansoni* infection status and intensity with periportal fibrosis irrespective of the diagnostic for schistosome infection. No support was found for current WHO elimination guidelines despite using arguably the most biologically specific and severe morbidity associated with schistosomiasis. We also

characterized the age-specific likelihood of periportal fibrosis, identifying a transitional age as young as 25 years. We identified future avenues for research into coinfections such as HIV and hepatitis B that appear to influence periportal fibrosis status even after controlling for a wide range of biosocial determinants of schistosome infection, treatment, and unrelated liver fibrosis. Future work is needed to understand if/how coinfections alter the pathogenesis of periportal fibrosis. Importantly, World Health Organization guidelines should be differentiated for schistosomiasis morbidities to discourage the use of infection status/intensity/prevalence as a proxy indicator for monitoring the elimination of periportal fibrosis as a public health problem.

INTRODUCTION

Fifty-two countries require mass drug administration (MDA) for schistosomiasis with over 250 million people estimated to be infected and over 700 million people living in endemic areas¹. The intestinal forms of chronic schistosomiasis, most prevalent in sub-Saharan Africa where *Schistosoma mansoni* is endemic, are characterised by periportal fibrosis (PPF).¹ Mature schistosome flukes live in the mesenteric venules; paired female flukes release approximately 300-400 eggs per day.² Only half of the eggs released successfully cross the intestinal mucosa to be excreted in stool. Eggs can become lodged in the gut wall or swept back into the vasculature of the liver and spleen. PPF is a chronic hepatic condition caused by an inflammatory response leading to granulomas and eventual fibrosis around parasite eggs in the portal veins and its segmental branches. The fibrosis can cause blockage of the main portal vein leading to portal hypertension and portosystemic collaterals such as oesophageal varices that when ruptured confer a high probability of death.³ Despite the clinical severity, there are no routine treatment strategies for individuals with PPF. The current World Health Organization (WHO) guidelines for schistosomiasis⁴ focus on current infection as a proxy indicator for assessing morbidities associated with schistosomiasis, including PPF. Because of this focus, MDA with praziquantel to treat current infections only is the mainstay of treatment for schistosomiasis-associated conditions. However, there is no consensus as to the relevance of current schistosome infections for PPF.⁵⁻⁸

The epidemiology of PPF is poorly understood. In the context of mass drug administration and natural fluke death, PPF can outlast current schistosome infections.⁶⁻⁹ Existing studies have focused on unrepresentative samples of limited populations often targeting only adults.^{5,6,10-12} Information is lacking on the age-specific PPF likelihood over a cross-section of the endemic population currently eligible for praziquantel treatment (aged 5+ years). There are assumptions that the relationship of age with PPF is linear,^{6,8,13} despite intestinal schistosome prevalence being nonlinearly associated with age. Importantly, schistosome infections mostly have been studied in isolation for their associations to clinical liver pathologies,^{6,8} neglecting coinfections such as malaria or chronic hepatitis B/C and co-occurring conditions such as alcoholism which may confound or modify PPF.

We investigated PPF outcomes in a wide cross-section of age (5-90 years) for 2834 individuals in 38 rural communities within Western and Eastern Uganda. The aim of this study was to establish the epidemiology of PPF and the role of current individual schistosome infection status/intensity. We examined PPF likelihood over age, the contribution of current *S. mansoni* infection status to PPF, and the key biological and social risk factors associated to PPF with a particular focus on coinfections and comorbidities.

Methods

Participant sampling

This study was nested in the SchistoTrack Cohort, using the baseline data collection completed in January-February of 2022 in Buliisa, Pakwach, and Mayuge Districts in Uganda. MDA with praziquantel for individuals aged ≥ 5 years was started in 2003 with our districts receiving ≥ 13 rounds of treatment before this study with the most recent round in December 2020. Programmatic treatment coverage was on average 79.43% (std. dev. 21.93%), 80.18% (std. dev. 23.02%), and 77.74% (std. dev. 13.19%) in Buliisa, Pakwach, and Mayuge Districts, respectively. A total of 1459 households were randomly sampled using local registers from 38 villages. One child aged 5-17 years and one adult aged 18+ years were selected by the household head/wife and invited for the clinical assessments, while information also was collected on every member aged 1+ years from the sampled households.

Further details are provided in Supplementary Text S1. Figure 1 shows the participants sampled and the final number of 2834 individuals from 1442 households henceforth used for analysis. The number of individuals was calculated to detect a minimum effect size of ~8% in an unevenly schistosome exposed/unexposed population (0.57) with a household design effect (1.136) and unbalanced strata/clusters at 97.5 power to account for multiple comparisons.

Outcome

For PPF diagnosis, point-of-care B-mode ultrasound was used to acquire image patterns as described in the Niamey Protocol.¹⁴ Sonographers regardless of experience with the Niamey Protocol were retrained as part of the study. Inter-observer agreement among the eight study sonographers was evaluated to reach a consensus on abdominal sweep procedures for acquiring each PPF pattern. At least two sonographers were required to jointly scan and agree on grading. All observed patterns of PPF were recorded. The highest pattern assigned to each participant was used to construct a binary outcome where one or any of the C-F combinations was coded as one (PPF) and patterns A-B were coded as zero (no PPF). Philips Lumify C5-2 curved linear array transducers were used with the Philips Lumify Ultrasound Application v3.0 on Lenovo 8505-F tablets with Android 9 Pie. Lossless DICOM images and videos were saved for quality assurance.

Exposures

S. mansoni infection was diagnosed using Kato-Katz (KK) microscopy¹⁵ and point-of-care circulating cathodic antigen tests (POC-CCA; batch 210811080).¹⁶ For KK, a single stool sample, double slide thick smears were prepared and read by two technicians with the results from the 41.7mg slides averaged and multiplied by 24 to calculate eggs per gram (EPG). KK infection status was positive when $EPG \geq 1$. WHO infection intensity categories were constructed with none, low, moderate, and heavy as 0, 1-99, 100-399, 400+ EPG respectively [8]. POC-CCA results were interpreted as negative, trace (barely visible test line), positive 1 (test line fainter than control line), positive 2 (test line similar to control line) and positive 3 (test line much darker than control line). Infection status by POC-CCA was positive if positive 1-3 was assigned; two alternative variables were constructed with trace positive (Tr+) and negative (Tr-). Ten percent re-readings of participant diagnoses by a senior technician were completed for KK slides and images of POC-CCA. Village-level variables were constructed from individual infection indicators (Supplementary Table S1).

Covariates

Detailed definitions are provided in Supplementary Text S1. At the individual level, we measured age, gender, religion, tribe, water contact, education, and occupation. The history of liver diseases, human immunodeficiency virus (HIV+) status, previous use of praziquantel or antiretroviral therapy (ART), complicated/uncomplicated malaria, hepatitis B/C, alcohol use, and smoking were recorded. Ultrasound-based first-pass assessments of ascites, chronic hepatitis/early cirrhosis-like disease, and fatty-like livers were completed, and rapid diagnostic tests of malaria were conducted (SD Bioline Malaria Ag P.f/Pan). At the household level, variables were constructed for total years of settlement in the village, home quality score, electricity availability, social status, home ownership, and improved drinking water, hygiene, and sanitation. Other variables included water body types in the village of residence, availability of at least one public latrine within the village, and district.

Statistical analyses

All the analyses were conducted with R v4.1. The functional form of the relationship between age and PPF was identified using the locally weighted scatterplot smoothing

(LOWESS) and Generalised Addictive Models (GAM) where spans were chosen through 10-fold cross-validation (CV).¹⁷ To identify key transition points in PPF likelihood across age, an alternative method of free-knot spline modelling framework with generalised CV was used.¹⁸

The intraclass correlation coefficients (ICC) in an empty multi-level logistic model was used to examine the maximum clustering of PPF within households and villages.¹⁹ A Hausman test²⁰ was used to compare the empty multilevel model with an empty logistic model to investigate village fixed versus random effects. PPF predictors were chosen with likelihood ratio tests (LRTs) of univariate models when p-values were ≤ 0.05 . ART and other variables with $\leq 1\%$ frequency was dropped (Table 1). To account for the sampling design, fully adjusted multivariable logistic models were run with clustered standard errors at the household level.²¹ Covariates selected by LRTs but with variance inflation factor (VIF) > 10 were removed and presented in alternative models. The model building procedure was rerun for subgroups of only children, adults, and individuals from Pakwach District. Participants with both PPF, history of liver diseases, chronic hepatitis/early cirrhosis-like disease, or fatty-like livers (Niamey Protocol XY patterns) were recoded as zero (no PPF) and models were rerun. To assess the predictive capacity of models, 10-fold CV was used with 5-fold CV for sub-group analyses with fewer observations.²²

Ethics approvals

Data collection and use were reviewed and approved by Oxford Tropical Research Ethics Committee (OxTREC 509-21), Vector Control Division Research Ethics Committee of the Uganda Ministry of Health (VCDREC146), and Uganda National Council of Science and Technology (UNCST HS 1664ES).

Role of funding source

The study sponsor had no role in the design, collection, analysis and interpretation of the data or the writing of the report.

RESULTS

PPF and *S. mansoni* prevalence

Participant characteristics are shown in Table 1 with liver fibrosis patterns demonstrated in Figure 2. PPF prevalence is shown in Figure 3a-d. Household and village covariates are presented in Supplementary Tables S2. Over 12% (344/2834) of participants were diagnosed with PPF; C image patterns were most common (5.39%, 153/2834). The highest prevalence of PPF at 19.46% was found in Pakwach (180/925), followed by 12.11% (115/950) in Buliisa and 4.90% (47/959) in Mayuge District. Across all individuals, PPF prevalence was higher among males when compared to females, and among adults when compared to children. Only 5 of 2834 (0.18%) participants presented with any grade of ascites; all ascites cases had PPF. There were 2.34% (69/2834) and 3.32% (94/2834) of participants, respectively, with chronic hepatitis/early cirrhosis-like or fatty-like livers detected by ultrasound. Over 7% (24/342) participants with PPF had a chronic hepatitis/early cirrhosis-like liver. By KK, 43.37% (1229/2834) of individuals were infected, with 8.22% (233/2834) heavily infected. POC-CCA Tr+ at 65.73% (1863/2834) was higher than KK prevalence and POC-CCA Tr- at 40.86% (1158/2834) was similar to KK prevalence.

Age-specific PPF

The study participants had a median age of 18 years (IQR 9-37) with a range of 5-90 years. Figure 3e-f presents the functional form of age with PPF. Age showed a linear relationship from childhood through to adolescence (5-25 years) with the likelihood of PPF changing exponentially in young adulthood to middle age (26-45 years). The likelihood of PPF becomes steady until around 55-60 years of age where there is a sharp decline in the likelihood of PPF approaching <10%. A transitional point in PPF likelihood was observed around age 25-26 years (Supplementary Figure S1). Both LOWESS and GAM analyses suggested non-monotonic relationships with an approximately quadratic relationship (inverse-U) of age with PPF.

Determinants of PPF

No intraclass clustering of PPF was observed within households, whereas 14% of variation was found within villages. Figures 4-5 show unadjusted associations. In fully adjusted models (Figure 6 with alternative infection indicators in Supplementary Figure S2), no individual-level *S. mansoni* infection indicators were associated with PPF. This result was robust to sub-group analyses for children, adults, Pakwach District only, and recoding of comorbid diffuse liver disease (XY patterns) as no PPF (Supplementary Figures S4-S7). Each one-year increase in age was associated with a 14.3% increase in the likelihood of PPF (95% CI 1.10-1.18). The quadratic term for age that captured the inverse-U relationship showed a negative association with PPF, although this effect was small <0.001% (95% CI 0.998-0.999). Female participants were 32.2% less likely than males to have PPF (95% CI 0.49-0.94). Fishermen were 71.2% more likely than unemployed or other occupations to have PPF (95% CI 1.12-2.63). Subsistence farmers were 32.6% less likely to have PPF when compared to unemployed or other occupations (0.47, 0.98). Participants with a reported previous diagnosis of liver diseases and HIV+ by a government health worker had 2.19 and 2.31-fold increased likelihood of PPF, respectively. Participants with ultrasound-detected chronic hepatitis/early cirrhosis-like disease had a 2.84-fold higher likelihood of PPF (95% CI 1.42-5.67) than individuals without these ultrasound-detected diffuse liver diseases. No household-level factors were significantly associated with PPF.

Village and district characteristics were associated with PPF (Figure 6). Living in a village with <5% prevalence of heavy infection intensity (PHI) among adults was borderline significantly associated with a 30.2% reduced likelihood of individual PPF (95% CI 0.48-1.00). Living in a village with ≥ 1 public latrine was associated with a 66.8% higher likelihood of individual PPF (95% CI 1.16-2.40). Only 1/3rd or less of the public latrines across the villages were reported to be in use at the time of the survey due to flooding. Participants living in Buliisa had 1.99 and in Pakwach had 2.81-fold higher likelihood of PPF than participants living in Mayuge District. When exact tribe was investigated (Supplementary Figure S3), participants from the Alur tribe were found to be 78.3% more likely to have PPF when compared to other tribes. Alur and Bagungu tribes were similar with respect to the percentage of fishermen (10.89%; 167/1534 vs. 10.97%; 26/237), whereas the Musoga tribe had fewer fishermen (4.33%; 20/462). The tribes of Alur, Bagungu, and Musoga were highly collinear with district effects (VIF>10). All models showed high predictive accuracy of PPF with 10-fold CV area under the receiver operating characteristic (ROC) curve of >80%.

DISCUSSION

There are no guidelines within the recent 2020-2030 WHO Roadmap for Neglected Tropical Diseases²³ for routine surveillance and treatment of morbidity associated with

schistosomiasis. Here we showed that severe morbidity caused by *S. mansoni* persisted in areas that have received annual community-wide MDA. We conducted a population-based study of 2834 individuals aged 5-90 years across 38 rural villages in diverse regions of Uganda. Our findings showed that age-specific PPF likelihood was nonlinear, drastically changing around age 25 years. Underlying liver diseases associated with the history of HIV coinfection and ultrasound-detected chronic hepatitis/early cirrhosis-like livers as opposed to current singular *S. mansoni* infections were key predictors of having PPF.

Age-specific PPF likelihood was nonlinear, suggesting the association of age with PPF changes over the life course of an individual. The trends found for PPF likelihood do not agree with well-established age-specific infection curves.²⁴ Age-specific infection prevalence and mean intensity are known to peak in school-aged children and decline sharply in adults, which has provided past evidence for acquired immunity even in the context of steady infection exposure.²⁵ Yet, PPF likelihood appears to linearly increase with age in children whilst exponentially increasing in young adults. There appeared to be a proximal transition age of 25 years in likelihood of PPF before levelling off in older adults and sharply declining in the eldest individuals. These results suggest that other factors unrelated to cumulative history of schistosome infection are co-occurring during early adulthood and influencing PPF onset or progression. The drastic decline in older age, beginning from age 55-60, may indicate excess mortality due to *S. mansoni*-associated morbidities in our study areas as the average life expectancy in rural Uganda is 65 years.²⁶ Future work is ongoing in SchistoTrack to identify prospectively the risk factors that individuals encounter over their life course, the onset of PPF, and the likelihood of death due to PPF.

There was no consistent association of individual *S. mansoni* infection status/intensity or village-level prevalence irrespective of diagnosis with KK microscopy or POC-CCA in either unadjusted or adjusted regressions. This lack of association was robust even when examining only adults. Furthermore, recent receipt of praziquantel was not correlated with PPF likelihood suggesting current treatment through MDA which has been shown to approximate past treatment likelihood²⁷ may not address severe *S. mansoni* morbidity. This finding agrees with a meta-analysis by Andrade et al. showing morbidities associated with *S. mansoni* have poorer resolution rates after MDA as compared to *S. haematobium*.²⁸ FibroSchot is an ongoing trial in Western Uganda that is causally assessing the potential of praziquantel outside of MDA to prevent and reverse PPF.²⁹ The long history of MDA (≥ 13 years) in our study areas coupled with high reported programmatic coverage ($>75\%$) in individuals aged ≥ 5 years and well-demonstrated poor cure rates with praziquantel further suggest that blanket treatment campaigns cannot geographically eliminate PPF.^{27,30}

Infection prevalence and intensities in school-aged children were poor proxy indicators of the likelihood of an individual—whether another child or adult—having PPF. However, the PHI in adults was suggestive of a borderline positive association with PPF. Future work might investigate revising WHO proxy indicators. Village infection prevalence, including groupings of WHO categories used to guide the frequency of mass drug administration, were not consistently associated with PPF likelihood.⁴ This result may indicate that current guidance for treating infections through MDA cannot be used for hotspot identification of likely areas to have residual PPF. Any level of village infection prevalence could correlate with PPF likelihood. However, the persistence of heavy intensity infections in only adults could be investigated in future work as an indicator of community transmission intensity and potentially used for PPF hotspot identification. Given that acquired immunity to reinfection

in adults for *S. mansoni* has been well demonstrated.²⁵ the persistence of heavy intensity infections in adults might reflect an excessively high historic force-of-infection.

PPF shares some epidemiology with *S. mansoni* infection yet has little overlap with predictors of praziquantel receipt. Age, gender, and being a fisherman were found to be significant predictors of PPF and in the same directions as expected for predicting infection likelihood.²⁴ Yet, the effect sizes were approximately 1/4th or less the size of the effects for HIV+ history, chronic-hepatitis/early cirrhosis-like disease, and study district. Being a fisherman is related to water contact and, in turn, infection or reinfection risk.²⁴ However, it is unclear whether being female confers some protection against PPF due to water contact behaviours or differences to health care access. Women have been shown to access health care more routinely within their reproductive years³¹ and signs of liver fibrosis may be picked up at an early stage. For example, any liver enlargement, tenderness, or changes in consistency might be found incidentally during abdominal palpations for pregnancy which then warrant further examinations of the liver and lead to care. Key praziquantel treatment predictors including social status, home quality (wealth), education, and home ownership, among other factors were not relevant for predicting PPF.²⁷ This finding further supports that both direct and indirect measurements of past praziquantel treatment show no correlation with PPF.

Commonly assessed co-occurring conditions such as alcohol use and malaria were not associated with PPF, despite correlations with PPF in other studies^{32,33} that did not adjust for diffuse liver diseases or HIV+ history. It is possible the correlation of alcoholism with PPF in adults was not detectable in our study if it manifested in older individuals with advanced cirrhosis where our data may have exhibited survivor bias associated with PPF.³⁴ The lack of association of current drinking with PPF was not due to accounting for chronic hepatitis/early cirrhosis-like disease; when this variable was removed from the model, current drinking remained insignificant. For the malaria insignificance for PPF, the parasite predominantly affects the spleen. The liver is only briefly involved in the life cycle where sporozoites mature into schizonts, especially for *P. falciparum* which was most common in our study areas. The peak of malaria prevalence and its associated severe morbidities often are amenable to antimalarials and found in children <5 years of age who were not studied here.

History of HIV+ and ultrasound-detected chronic hepatitis/early cirrhosis-like disease were positively correlated with PPF. Each comorbidity has been shown previously to singularly contribute to the likelihood of PPF.^{7,35} HIV+ has been shown to correlate with hyaluronic acid levels,³⁶ which are indicative of liver fibrosis. What remains unclear is the mechanism of HIV+ and *S. mansoni* interactions on the ability of the host to mount a granulomatous response to eggs or for the associations with prognostic outcomes such as severe anaemia or portal hypertension. For HIV+, it is possible that status was underestimated if there were individuals who have yet to be diagnosed or who had concerns in reporting due to community stigmatisation. Chronic hepatitis or cirrhosis causes diffuse liver pathologies affecting liver functionality. Confirmatory testing that considers viral hepatitis diagnostics, alternative imaging modalities, and fibrosis scores is needed in future work for distinguishing and staging hepatitis from cirrhosis. Considering causes of hepatitis-like livers, hepatitis C is rare in our study areas, whereas hepatitis B may be a promising line of investigation as it has been shown to exceed 10% in adults in Western Uganda,³⁷ and will be incorporated in future SchistoTrack follow-ups. Confounding by aflatoxins also must be investigated, which is challenging given the diverse subtypes and sampling frames of either human or grain. Importantly, for future investigations of diffuse liver disease interactions with PPF, an

important consideration is to separate image pattern outcomes. A PPF pattern of D is less diffuse than other PPF patterns; hence, viral hepatitis coinfections might be suspected to only correlate with specific image patterns.

We found that individuals in Pakwach and Buliisa Districts were more likely to have PPF than individuals in Mayuge District. Geographical differences in PPF have been documented previously between Western and Eastern Uganda.^{13,38} Our results suggest the district effects are not due to environmental conditions and, in turn, possible differences in parasite strains. No ecological predictors were correlated with PPF. Moreover, individual tribes were highly collinear with district effects. Although there was no household clustering of PPF, we are unable to make any inferences on the relevance of genetic ties for PPF as we did not require the sampled adult to be the parent of the sampled child. To identify any shared genetic effects through tribe, there is a need to prospectively examine only child-adult or sibling pairs. Research is ongoing in SchistoTrack to assess the development of PPF across a five-year period in children belonging to households that have parents with PPF. Alternatively, the effect of tribe might represent differences in water contact behaviours not captured through differences in the prevalence of fishermen or differences in access to health care.^{39,40}

Indicators for current individual infection status/intensity were not associated with the likelihood of PPF. Proxy indicators for monitoring morbidity should be tailored to the epidemiology of the morbidity. It is well-established that PPF is more prevalent in adults. PHI in adults, rather than the current WHO guideline noting PHI in school-aged children, might be a better indicator to assess progress towards control of morbidities associated with *S. mansoni*. Programmes aimed at controlling PPF should consider an integrated approach, especially for adults in endemic settings. Improved access to health care services for the treatment of schistosomiasis and capacity building within local health systems for recognising PPF are needed. Multimorbidity could be addressed through case management strategies to explore early identification of PPF. Areas where integration across diseases might be explored may include investigating how hepatitis B/C and HIV clinics could be expanded to include diagnosis for schistosomiasis-associated morbidity. The results of this study, if replicated in other settings, could provide interventions to be tested in future randomised-controlled trials for improving morbidity management for severe hepatic schistosomiasis.

Author contributions

Conceptualization: GFC and NBK. Data curation and validation: SA. Formal analysis: SA. Investigation, methodology, validation, visualization, and writing – original draft: SA and GFC. Writing – review and editing: SA, BN, SM, BK, CKO, NBK, and GFC. Funding acquisition and supervision: GFC. Resources: BN, GFC, and NBK. Data collection: BN, SM, BK, NBK, and GFC.

Data sharing

Data is not publicly available due to data protection and ethics restrictions related to the ongoing nature of the SchistoTrack Cohort and easily identifiable nature of the participants.

Declarations of interest

BN, SM, BT, and NBK were or are employed by the Uganda Ministry of Health. All other authors declare no conflicts of interest.

Acknowledgements

This work was supported by the Nuffield Department of Population Health Pump Priming Fund, Wellcome Trust Institutional Strategic Support Fund (204826/Z/16/Z), John Fell Fund, Robertson Foundation Fellowship, and UKRI EPSRC Award (EP/X021793/1) to G.F. Chami. We are thankful for the involvement and trust from our study participants. There were extensive SchistoTrack field teams involved including malacologists, surveyors, nurses, sonographers, and laboratory technicians. Without the support of the Uganda Ministry of Health, local district leaders, focal health workers, and village health teams, we could not have worked so closely in partnership with the study communities. We also have many thanks to give to the Oxford team for support with fieldwork, data wrangling, and feedback during group meetings.

Table 1 Participant characteristics

Variable	Overall N = 2834 n (%)	No PPF N = 2492 N (%)	PPF N = 342 N (%)
≥ 1 EPG	1229 (43·37)	1089 (43·70)	140 (40·94)
Infection Intensity (KK)			
No	1605 (56·63)	1403 (56·30)	202 (59·06)
Low	632 (22·30)	563 (22·59)	69 (20·18)
Mild	364 (12·84)	320 (12·84)	44 (12·87)
High	233 (8·22)	206 (8·27)	27 (7·89)
EPG, median (IQR)	0 (0-72)	0 (0-72)	0 (0-60)
LN (EPG + 1), median (IQR)	0 (0- 4·29)	0 (0- 4·29)	0 (0-4·11)
POC-CCA readings			
Negative	971 (34·26)	826 (33·15)	145 (42·40)
Trace	705 (24·88)	630 (25·28)	75 (21·93)
Pos.1	543 (19·16)	497 (19·94)	46 (13·45)
Pos.2	438 (15·46)	386 (15·49)	52 (15·20)
Pos.3	177 (6·25)	153 (6·14)	24 (7·02)
Infection status (POC-CCA, Tr-)	1158 (40·86)	1036 (41·57)	122 (35·67)
Infection status (POC-CCA, Tr+)	1863 (65·74)	1666 (66·85)	197 (57·60)
History of liver disease	72 (2·54)	49 (1·97)	23 (6·73)
History of HIV+	44 (1·55)	31 (1·24)	13 (3·80)
Received anti-retroviral therapy (ART)	17 (0·60)	11 (0·44)	6 (1·75)
History of hepatitis B/C	25 (0·88)	21 (0·84)	4 (1·17)
PfHRP2	937 (33·06)	862 (34·59)	75 (21·93)
PfHRP2/pLDH (invalid -)	1119 (39·48)	1038 (41·65)	81 (23·68)
Malaria (invalid-)			
Negative	1715 (60·52)	1454 (58·35)	261 (76·32)
PfHRP2 & pLDH/pLDH only	178 (6·28)	173 (6·94)	5 (1·46)
PfHRP2 only	941 (33·20)	865 (34·71)	76 (22·22)
Praziquantel within past 1 year	1341 (47·32)	1158 (46·47)	183 (53·51)
Early cirrhosis-like (US)	5 (0·18)	3 (0·12)	2 (0·58)
Fatty-like (US)	94 (3·32)	81 (3·25)	13 (3·80)
Chronic hepatitis-like (US)	64 (2·26)	42 (1·69)	22 (6·43)
Chronic hepatitis/Early cirrhosis-like (US)	67 (2·36)	43 (1·72)	24 (7·02)
Age	18 (9·37)	15 (9·34)	39·50 (30-50)
Religion			
Christian	2361 (83·31)	2054 (82·42)	307 (89·77)
Muslim	456 (16·09)	425 (17·05)	31 (9·06)
no/others	17 (0·60)	13 (0·52)	4 (1·17)
Majority religion	2057 (72·58)	1789 (71·79)	268 (78·36)
Majority tribe	2080 (73·39)	1797 (72·11)	283 (82·75)
Tribes			
Other	601 (21·21)	558 (22·39)	43 (12·57)
Musoga	462 (16·30)	443 (17·78)	19 (5·56)
Bagungu	237 (8·36)	207 (8·31)	30 (8·77)
Alur	1534 (54·13)	1284 (51·52)	250 (73·10)
Educational attainment (years), median (IQR)	3 (1, 5)	3 (1, 5)	4 (2, 6)
Occupation			
None/others	1978 (69·80)	1819 (72·99)	159 (46·49)
Subsistence farmer	500 (17·64)	430 (17·26)	70 (20·47)
Fisherman	248 (8·75)	154 (6·18)	94 (27·49)
Fishmonger	108 (3·81)	89 (3·57)	19 (5·56)
Current drinker	235 (11·17)	165 (9·30)	70 (21·28)
Current smoker	157 (7·46)	110 (6·20)	47 (14·29)
Any current water contact	1327 (46·82)	1109 (44·50)	218 (63·74)

PfHRP2 = *Plasmodium falciparum* antigen histidine rich protein 2. pLDH = parasite lactate dehydrogenase (generic to all malaria species).
For pLDH only, there were only four cases. US = Ultrasound

FIGURE LEGENDS

Figure 1 Participant sampling. The selection of study participants is shown and more fully described in the Methods section.

Figure 2 Periportal fibrosis image patterns and comorbidities. Patterns as described in the Niamey Protocol are shown and suspected diffuse liver diseases such as chronic hepatitis/early cirrhosis-like livers or fatty-like livers are noted where observed. **a.** A normal liver pattern is shown with no comorbidities. **b.** A B1 image pattern depicting flying saucers or starry sky is shown. **c.** A B2 spider thickening pattern is shown. **d.** A C1 image pattern (prominent peripheral rings) with a chronic hepatitis-like liver is shown. The diffuse disease was suspected to be more hepatitis-like in pattern than cirrhosis-like due to the retained liver surface regularity. **e.** A C1 image pattern is shown with a more cirrhosis-like, possibly advanced cirrhosis, comorbidity where there was evidence of a possible shrunken liver and gross surface irregularity. **f.** A C2 (prominent pipe stems) image pattern is shown. **g.** A D image pattern (ruff portal bifurcation) is shown. **h.** Image Pattern E (patches representing occluded, bright white vessels) is shown. **i.** The most severe pattern of PPF (F, a bird's claw) is shown. **j.** A fatty-like liver with no evidence of image patterns B-F is shown.

Figure 3 Prevalence of periportal fibrosis and age-specific likelihood. N=2834. Pakwach N = 925. Buliisa N = 950, Mayuge N= 959. **a.** The prevalence of liver patterns is shown, whereby only the highest pattern per participant was included. **b.** Prevalence of periportal fibrosis (liver pattern C-F) by study districts and age group is shown. **c.** Prevalence of periportal fibrosis by study district and gender is shown. **d.** Prevalence of *S. mansoni* infection by periportal fibrosis is shown. **e.** Locally weighted scatterplot smoothing (LOWESS) was applied. The panel shows fewer data points among the study sample between the ages of 16-19 years. The predicted probability of periportal fibrosis at each value of age using the LOWESS default span (0.75) is shown. **f.** Similar to panel e, the predicted probability of periportal fibrosis is shown at each value of age, but differently here using a span of 0.62 selected by 10-fold CV.

Figure 2 Unadjusted models for *S. mansoni* infections and periportal fibrosis. Odds ratios (ORs) and corresponding 95% confidence interval (CI) represent univariate logistic regression models. **a.** Unadjusted relationships are shown between periportal fibrosis and sociodemographic indicators. **b.** Unadjusted relationships are shown between periportal fibrosis and morbidity indicators.

Figure 3 Unadjusted models for covariates and periportal fibrosis. Odds ratios (ORs) and corresponding 95% confidence interval (CI) represent univariate logistic regression models. PfHRP2 = *Plasmodium falciparum* antigen histidine rich protein 2. pLDH = parasite lactase dehydrogenase (generic to all malaria species). For pLDH only, there were only four cases. US = Ultrasound. **a.** Unadjusted relationships are shown between PPF and coinfections as well as other medical history factors. Alcohol and smoking variables were run for adult only models (individuals aged 18+ years). **b.** Unadjusted relationships are shown between PPF and sociodemographic and ecological indicators.

Figure 4 Determinants of periportal fibrosis. A fully adjusted model is shown with variables selected through LRT with p-value ≤ 0.05 . Black represents significant relationships of p-value ≤ 0.05 . No variables in the model had a VIF > 10 . The mean area under the ROC (by 10-fold CV) was 0.82. Hausman test result of an empty random vs village fixed effects

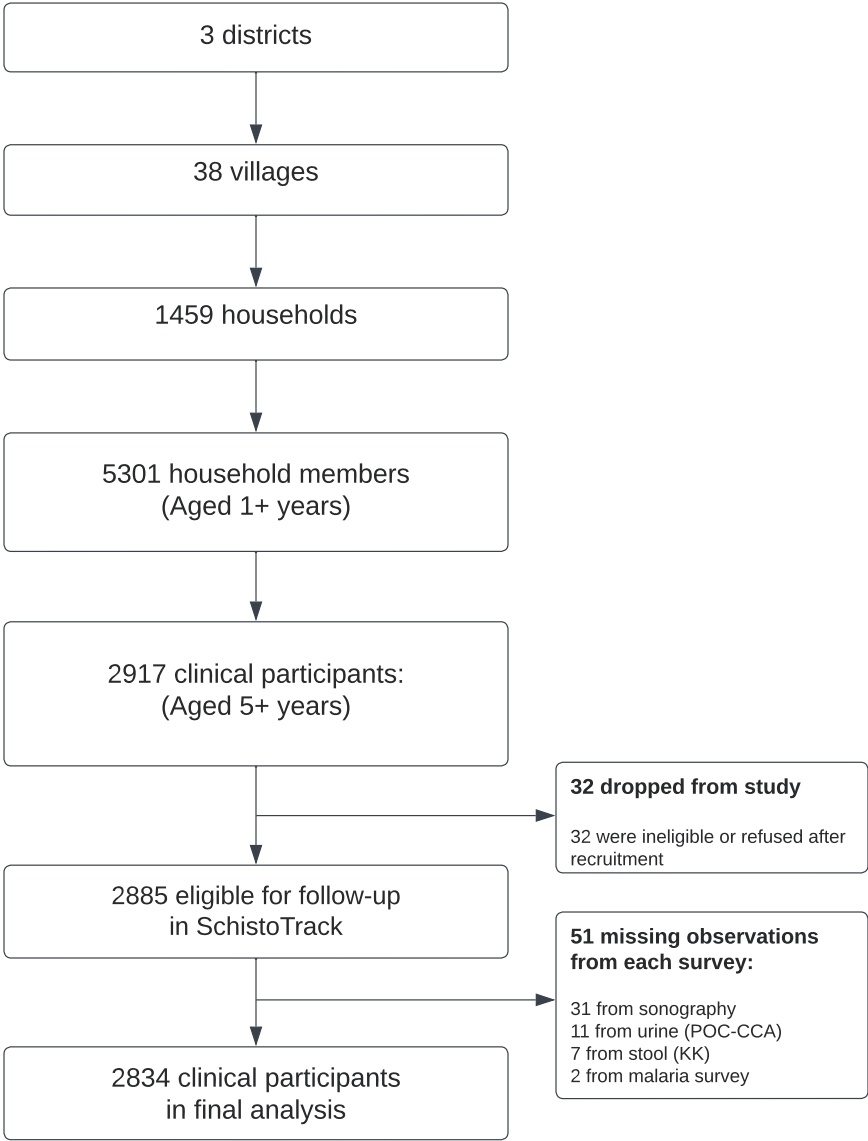
model (with 38 villages) was $\chi^2 = 3.90$, $p\text{-value} < 0.05$. PfHRP2 = *Plasmodium falciparum* antigen histidine rich protein 2. pLDH = parasite lactate dehydrogenase (generic to all malaria species). For pLDH only, there were only four cases. US = Ultrasound.

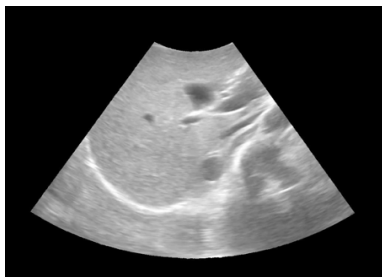
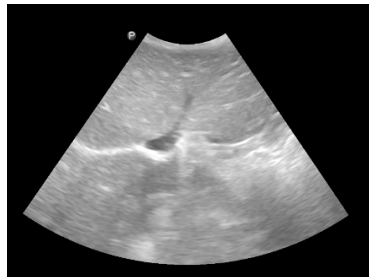
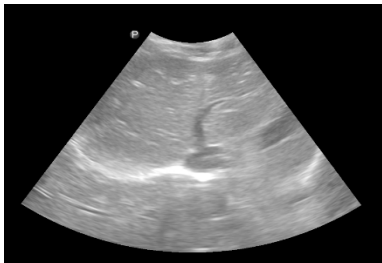
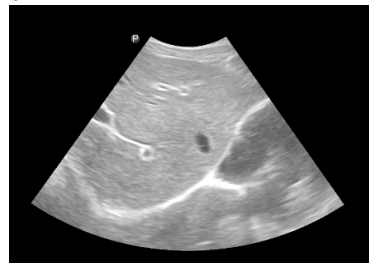
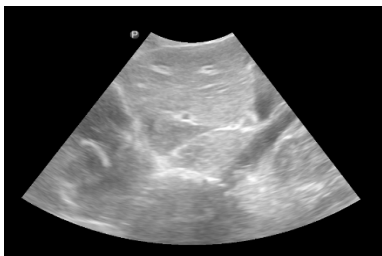
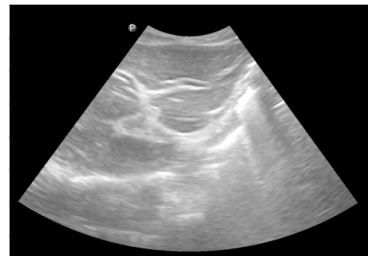
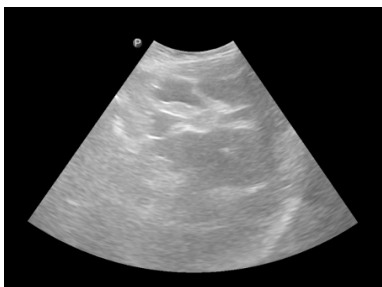
REFERENCES

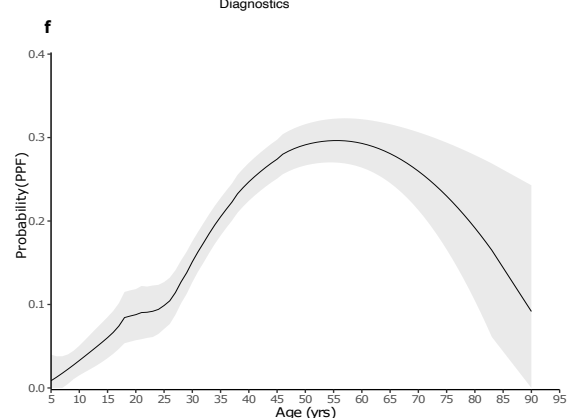
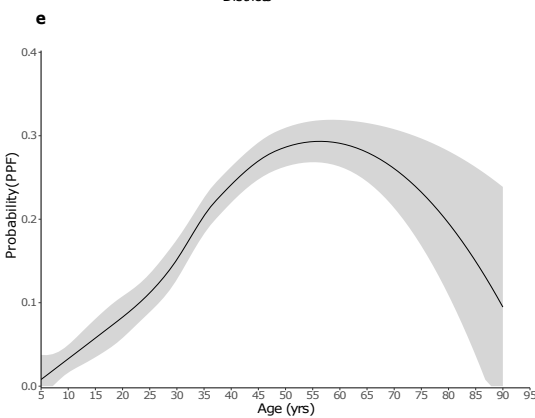
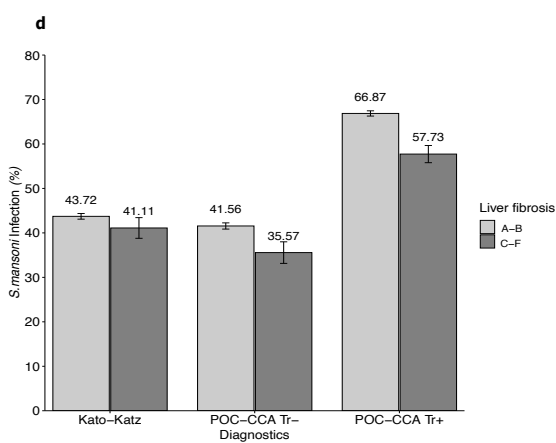
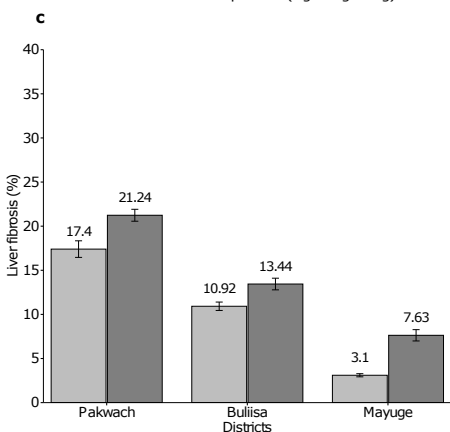
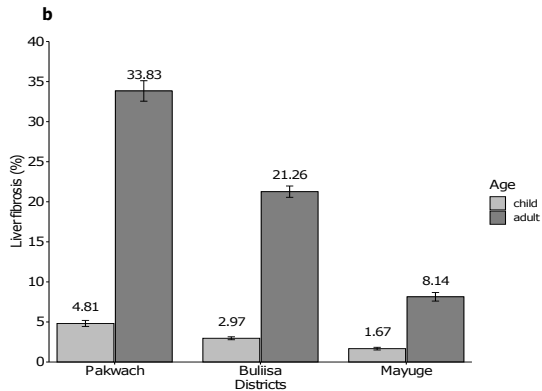
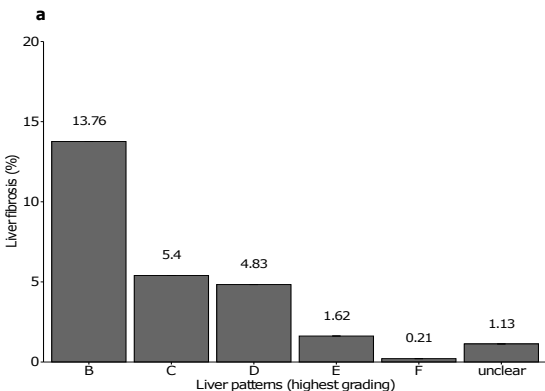
1. McManus DP, Dunne DW, Sacko M, Utzinger J, Vennervald BJ, Zhou XN. Schistosomiasis. Nat. Rev. Dis. Primers **2018**; 4:1–19.
2. Cheever AW, Macedonia JG, Mosimann JE, Cheever EA. Kinetics of egg production and egg excretion by *Schistosoma mansoni* and *S. japonicum* in mice infected with a single pair of worms. Am. J. Trop. Med. **1994**; 50:281–295.
3. Opio CK, Kazibwe F, Rejani L, Kabatereine NB, Ocama P. Hepatic schistosomiasis, upper gastrointestinal bleeding, and health related quality of life measurements from the Albert Nile Basin. J. Patient-Rep. Outcomes **2021**; 5.
4. WHO. WHO guideline on control and elimination of human schistosomiasis. Geneva: 2022. Available at: <https://www.who.int/publications/i/item/9789240041608>. Accessed 3 March, 2023.
5. Davis SM, Wiegand RE, Mulama F, et al. Morbidity associated with schistosomiasis before and after treatment in young children in Rusinga Island, Western Kenya. Am. J. Trop. Med. **2015**; 92:952–958.
6. Russell HJ, Penney JMS, Linder C, et al. A cross-sectional study of periportal fibrosis and *Schistosoma mansoni* infection among school-aged children in a hard-to-reach area of Madagascar. Trans. R. Soc. Trop. Med. Hyg. **2020**; 114:315–322.
7. Berhe N, Myrvang B, Gundersen SG. Intensity of *Schistosoma mansoni*, hepatitis B, age, and sex predict levels of hepatic periportal thickening/fibrosis (PPT/F): A large-scale community-based study in Ethiopia. Am. J. Trop. Med. **2007**; 77:1079–1086.
8. Nalugwa A, Nuwaha F, Tukahebwa EM, Olsen A. *Schistosoma mansoni*-associated morbidity among preschool-aged children along the shores of Lake Victoria in Uganda. Trop. Med. Infect. Dis. **2017**; 2.
9. Mueller A, Fuss A, Ziegler U, Kaatano GM, Mazigo HD. Intestinal schistosomiasis of Ijinga Island, north-western Tanzania: prevalence, intensity of infection, hepatosplenic morbidities and their associated factors. BMC Infect. Dis. **2019**; 19:1–12.
10. Mazigo HD, Dunne DW, Morona D, et al. Periportal fibrosis, liver and spleen sizes among *S. mansoni* mono or co-infected individuals with human immunodeficiency virus-1 in fishing villages along Lake Victoria shores, North-Western, Tanzania. Parasit. Vectors. **2015**; 8:1–13.
11. Kariuki HC, Mbugua G, Magak P, et al. Prevalence and familial aggregation of schistosomal liver morbidity in Kenya: evaluation by new ultrasound criteria. J. Infect. Dis. **2001**; 183:960–966.
12. Wiegand RE, Secor WE, Fleming FM, et al. Control and elimination of schistosomiasis as a public health problem: thresholds fail to differentiate Schistosomiasis morbidity prevalence in children. Open Forum Infect. Dis. **2021**; 8:1–10.
13. Booth M, Vennervald BJ, Kabatereine NB, et al. Hepatosplenic morbidity in two neighbouring communities in Uganda with high levels of *Schistosoma mansoni* infection but very different durations of residence. Trans. R. Soc. Trop. Med. Hyg. **2004**; 98:125–136.
14. Richter J, Campange G, Hatz G, Berquist NR, Jenkins JM. Ultrasound in schistosomiasis a practical guide to the standardized use of ultrasonography for the assessment of schistosomiasis-related morbidity. World Health Organization **2000**; :55. Available at: <https://apps.who.int/iris/handle/10665/66535>.
15. Katz N, Chaves A, Pellegrino J. A simple device for quantitative stool thick-smear technique in *Schistosomiasis mansoni*. Rev. Inst. Med. Trop. **1972**; 14:397–400.

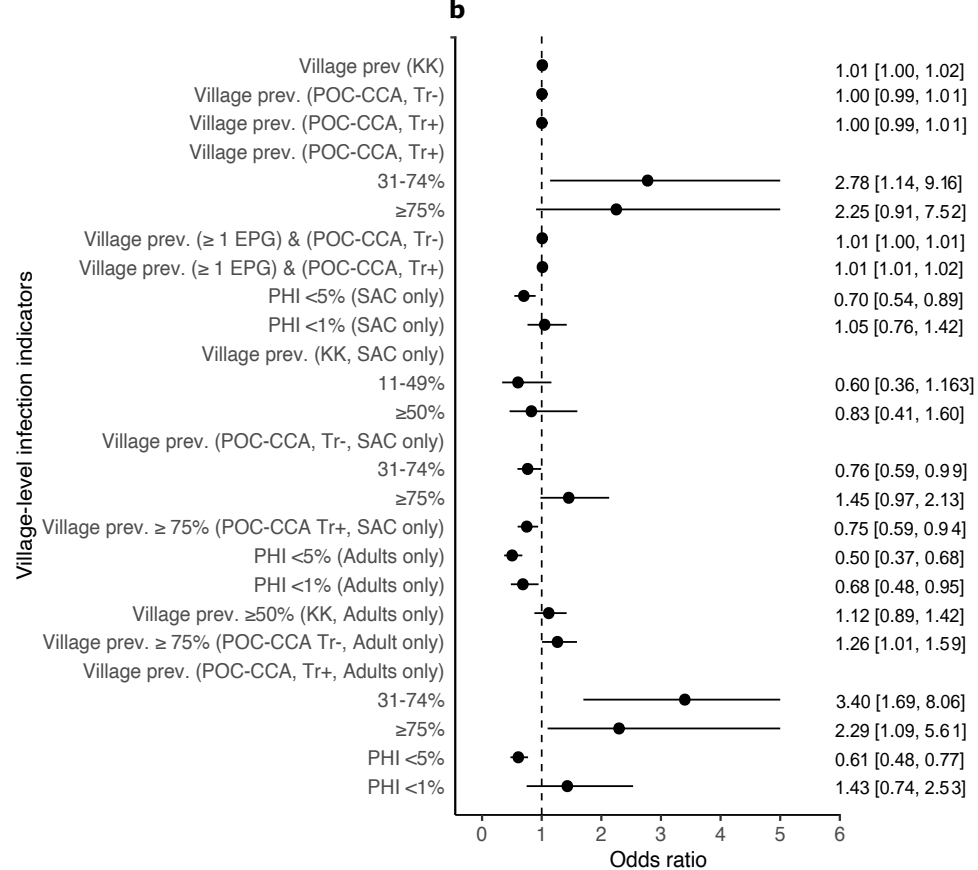
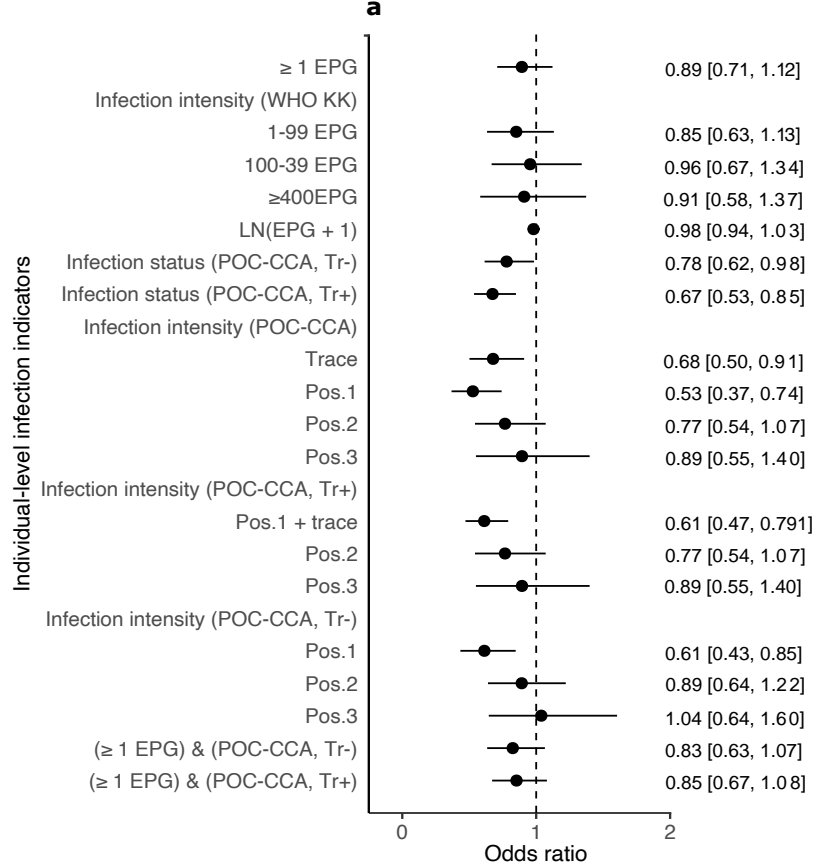
16. Casacuberta-Partal M, Hoekstra PT, Kornelis D, van Lieshout L, van Dam GJ. An innovative and user-friendly scoring system for standardised quantitative interpretation of the urine-based point-of-care strip test (POC-CCA) for the diagnosis of intestinal schistosomiasis: a proof-of-concept study. *Acta Trop.* **2019**; 199:105150.
17. Cleveland WS. Robust locally weighted regression and smoothing scatterplots. *J Am Stat Assoc.* **1979**; 74:829–836.
18. Spiriti S, Philip Smith A, Lecuyer P. freeknotfit: Fit Free-Knot Splines To Data. Algorithms for Implementing Free-Knot Splines. 2019. Available at: <https://rdrr.io/cran/freeknotsplines/man/freeknotfit.html>. Accessed 17 May 2023.
19. Merlo J, Chaix B, Ohlsson H, et al. A brief conceptual tutorial of multilevel analysis in social epidemiology: Using measures of clustering in multilevel logistic regression to investigate contextual phenomena. *J Epidemiol Community Health.* **2006**; 60:290–297.
20. Hausman J. Soecification Test in Econometrics. *Econometrica* **1978**; 46:1251–1271.
21. Williams RL. A note on robust variance estimation for cluster-correlated data. *Biometrics* **2000**; 56:645–646.
22. Arlot S, Celisse A. A survey of cross-validation procedures for model selection. *Stat. Surv.* **2010**; 4:40–79.
23. WHO. Ending the neglect to attain the sustainable development goals: a road map for neglected tropical diseases 2021–2030. Geneva: World Health Organization Geneva, 2020. Available at: <https://www.who.int/publications/i/item/9789240010352>. Accessed 8 April 2023
24. Chami GF, Fenwick A, Bulte E, et al. Influence of *Schistosoma mansoni* and Hookworm Infection Intensities on Anaemia in Ugandan Villages. *PLOS Negl. Trop. Dis.* **2015**; 9:1–18.
25. Kabatereine NB, Vennervald BJ, Ouma JH, et al. Adult resistance to *schistosomiasis mansoni*: age-dependence of reinfection remains constant in communities with diverse exposure patterns. *Parasitology* **1999**; 118:101–105.
26. WHO. Uganda: Health data overview for the republic of Uganda. 2023. Available at: <https://data.who.int/countries/800>. Accessed 19 June, 2023
27. Chami GF, Kontoleon AA, Bulte E, et al. Profiling Nonrecipients of Mass Drug Administration for Schistosomiasis and Hookworm Infections: A Comprehensive Analysis of Praziquantel and Albendazole Coverage in Community-Directed Treatment in Uganda. *Clin. Infect. Dis.* **2016**; 62:200–207.
28. Andrade G, Bertsch DJ, Gazzinelli A, King CH. Decline in infection-related morbidities following drug-mediated reductions in the intensity of *Schistosoma infection*: A systematic review and meta-analysis. *PLOS Negl. Trop. Dis.* **2017**; 11:1–23.
29. Mawa PA, Kincaid-Smith J, Tukahebwa EM, Webster JP, Wilson S. Schistosomiasis morbidity hotspots: roles of the human host, the parasite and their interface in the development of severe morbidity. *Front. Immunol.* **2021**; 12:1–21.
30. Zwang J, Olliaro PL. Clinical efficacy and tolerability of Praziquantel for intestinal and urinary schistosomiasis—a meta-analysis of comparative and non-comparative clinical trials. *PLOS Negl. Trop. Dis.* **2014**; 8.
31. Sserwanja Q, Mukunya D, Musaba MW, Kawuki J, Kitutu FE. Factors associated with health facility utilization during childbirth among 15 to 49-year-old women in Uganda: evidence from the Uganda demographic health survey 2016. *BMC Health Serv. Res.* **2021**; 21:1–13.
32. Wilson S, Jones FM, Mwatha JK, et al. Hepatosplenomegaly associated with chronic malaria exposure: Evidence for a pro-inflammatory mechanism exacerbated by

- schistosomiasis. *Parasite Immunol.* **2009**; 31:64–71.
33. Malenganisho WLM, Magnussen P, Friis H, et al. *Schistosoma mansoni* morbidity among adults in two villages along Lake Victoria shores in Mwanza District, Tanzania. *Trans. R. Soc. Trop. Med. Hyg.* **2008**; 102:532–541.
 34. Rashighi M, Harris JE, Omer SM Hasan, et al. Alcohol consumption and risk of liver cirrhosis: a systematic review and meta-analysis. *Am J Gastroenterol.* **2017**; 176:139–148.
 35. Sereti I, Krebs SJ, Phanuphak N, et al. Persistent, Albeit reduced, chronic inflammation in persons starting antiretroviral therapy in acute HIV infection. *Clin. Infect. Dis.* **2017**; 64:124–131.
 36. Gunda DW, Mtui EF, Kilonzo SB, Kidenya BR, Maziago HD. *Schistosoma mansoni*-Hepatitis B co-infection among adult patients with periportal fibrosis: a cross sectional study. *Clin Microbiol Infect* **2020**; 10:136–143.
 37. Bwogi J, Braka F, Makumbi I, et al. Hepatitis B infection is highly endemic in Uganda: findings from a national serosurvey. *African Health Sciences* **2009**; 9:98–108.
 38. Tukahebwa EM, Magnussen P, Madsen H, et al. A very high infection intensity of *Schistosoma mansoni* in a Ugandan Lake Victoria fishing community is required for association with highly prevalent organ related morbidity. *PLOS Negl. Trop. Dis.* **2013**; 7; e2268
 39. De Moira AP, Fulford AJC, Kabatereine NB, et al. Microgeographical and tribal variations in water contact and *Schistosoma mansoni* exposure within a Ugandan fishing community. *Trop. Med. Int. Health.* **2007**; 12:724–735.
 40. Dowhaniuk N. Exploring country-wide equitable government health care facility access in Uganda. *Int. J. Equity Health.* **2021**; 20:1–19.



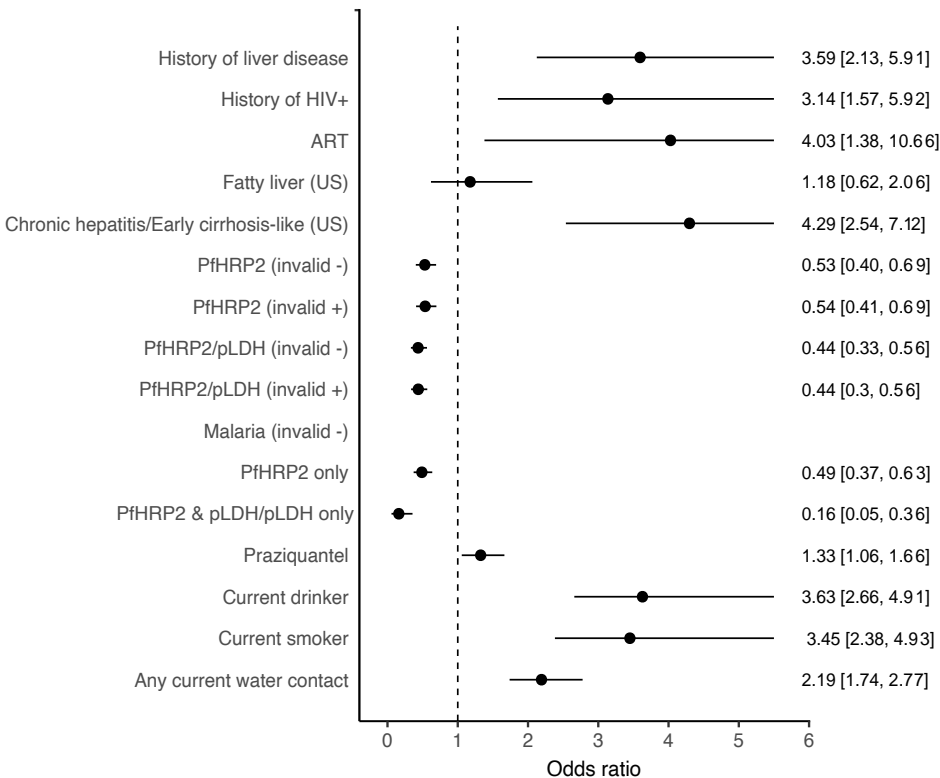
a**b****c****d****e****f****g****h****i****j**



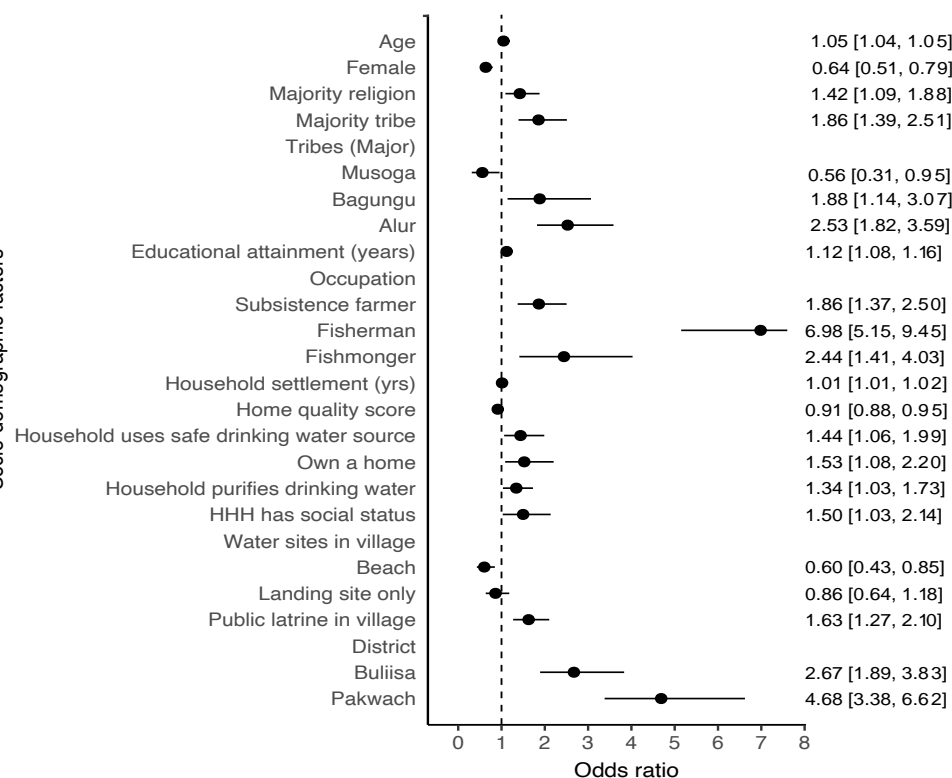


a

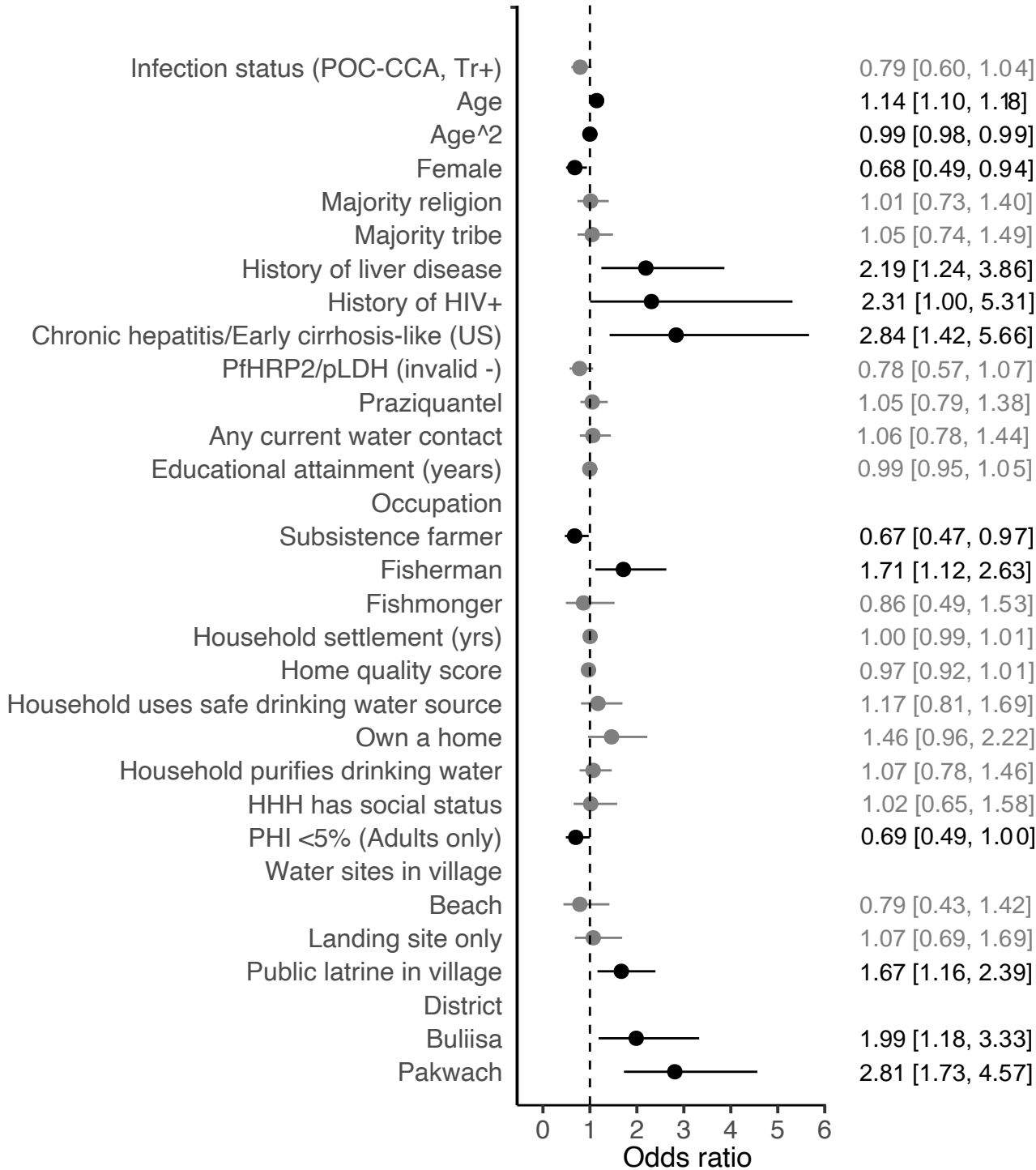
Medical history & behavioral factors

**b**

Socio-demographic factors



Factors



sig

● not-significant
● significant

Supplementary information for “The epidemiology of periportal fibrosis and relevance of current *Schistosoma mansoni* infection: a population-based, cross-sectional study”

Table of Contents

Text S1 Variable definitions	2
Table S1 <i>S. mansoni</i> infection indicators and their alternative definitions.....	4
Table S2 Descriptive statistics of household and village-level variables	5
Table S3 Descriptive statistics of villages-level infection indicators	6
Figure S1 Generalised Additive Model (GAM) with fitted value (logit) of PPF over age.....	7
Figure S2 Fully adjusted models with alternative infection indicators.....	8
Figure S3 Fully adjusted models with alternative comorbidities and sociodemographic factors.....	9
Figure S4 Subgroup analysis of participants from Pakwach District	10
Figure S5 Subgroup analysis of adult participants.....	11
Figure S6 Subgroup analysis of children participants.	12
Figure S7 Sensitivity analysis of XY liver patterns.....	13
References.....	14

Text S1 Variable definitions

Individual-level variables

Age was a continuous variable to the nearest year. Age squared was incorporated in models to account for nonlinearity with PPF. Gender and religion were defined as binary variables with females and Muslims coded as one, respectively. The majority religion was a binary variable, which was coded as one if participants belonged to the majority religion within their own village.

Medical history was recorded by household surveyors who used modified versions of the World Health Surveys to collect information on past diagnoses.¹ Individuals were asked whether a government health worker had diagnosed them with a set of conditions. Each condition was constructed as a binary variable. The conditions asked about included liver diseases (cirrhosis, scarring or fibrosis), HIV, and hepatitis B/C. The WHO Stepwise approach for risk factor surveillance (STEPS) was used to gather information on non-communicable diseases such as smoking and alcohol drinkers.² Individuals were coded as current smokers if any tobacco products were used within the past 12 months preceding the study. Individuals were coded as current drinkers if they had consumed any alcohol within the past 12 months. Only study participants aged 10+ years were surveyed on alcohol and smoking behaviours; therefore, these variables were only used in the adult sub-group model.

Previous treatment with praziquantel was coded as a binary variable where participants who received or ingested praziquantel in the past one year were coded as one. Mass drug administration began in all study districts in 2003. Buliisa and Pakwach Districts received 13 annual rounds of mass drug administration, whereas Mayuge District had received 15 rounds of mass drug administration preceding the SchistoTrack baseline. There was no mass drug administration the year before the survey due to COVID-19 lockdowns, so receipt of praziquantel would have been through other means such as leftover medicines remaining with community medicine distributors. The last round of mass drug administration in all our study districts was conducted in December 2020, using both school-based and community-based distribution. Current anti-retroviral therapy was also coded as one if participants confirm they have taken western medicine in the past one month and anti-retroviral therapy was indicated as one of them.

Ultrasound was used to gather information on suspected current diffuse liver diseases. Chronic hepatitis-like, cirrhosis-like, and fatty-like livers were noted.³ Chronic hepatitis-like livers were observed with diffuse, inhomogeneous echogenic nodules and no surface (liver outline) irregularities. Cirrhosis-like livers were defined as having a coarse liver parenchymal echotexture, rounded/blunted caudal liver edge, and potential surface irregularities. Possible shrunken livers or reduced left and right medial sections were considered albeit not required for our definitions of cirrhosis-like livers as we did not only record advanced cases. As chronic hepatitis can mimic early/compensated cirrhosis when diagnosed only via ultrasound,⁴ a binary variable equal to one was generated if there was any evidence of either chronic hepatitis-like or cirrhosis-like livers. In our data, all five cases of cirrhosis-like livers had sharp caudal edges suggesting early cirrhosis. Only one of five cases of cirrhosis-like livers had any (here, slightly serrated) surface irregularities. As we were interested in chronic underlying liver diseases, separate ‘starry-like’ patterns for acute hepatitis were not investigated outside the image pattern B provided in the Niamey Protocol. Fatty-like livers were defined as having diffuse parenchymal brightness with observed liver-to-kidney contrast and posterior attenuation. B-mode ultrasound has shown to be reliable for steatosis exceeding 20%,⁵ hence we were likely to miss mild fatty-like liver cases. Ascites was noted when fluid build-up was observed in any of the four abdominal quadrants. If fluid was observed, but not present in all four quadrants then it was graded as mild. If fluid was observed in all four quadrants, but the abdomen and organs were still clearly distinguishable then ascites was graded as moderate. When there was fluid build-up in all four quadrants to the point that the quadrants were indistinguishable, and organs were difficult to differentiate then ascites was graded as severe.

Any water contact was coded as one if a participant did at least one of following activities in the lake or river near their village including fishing, fishmongering, collecting papyrus/shells, retrieving drinking water, washing clothes with/without soap, washing jerry cans or other household supplies, swimming, or playing in water. Education was defined as a count variable from 1-14, whereas 1-7 represented primary school years, 8-13 captured lower and upper secondary school and 14 represented participants that have completed a post-secondary education. Education also was expressed as a categorical variable of no education, primary education, and secondary education or higher. Occupation was defined as a categorical variable of subsistence farmer, fishermen, fishmongers and no and other occupations as reference category.

Household-level variables including water, sanitation, and hygiene indicators, social status, home quality, electricity, ownership, and number of rooms as well as village ecology and infrastructure were defined as described in Chami et al. 2016.⁶ The years the household had settled in the village was a continuous variable defined as the years anyone in the household or an immediate relative of the household head (who was in the same household) had been living in the village. Outliers/typos in this variable were recoded to the next nearest value. The recoded outliers were -12 and 3333501, which were recoded to 12 and 50, respectively.

Table S1 *S. mansoni* infection indicators and their alternative definitions

Infection Indicator	Definition
LN(EPG + 1)	Natural log of average egg per gram of stool (+1)
Infection intensity (POC-CCA, Tr+)	Infection categories from urine sample using point-of-care circulating cathodic antigen (POC-CCA) test; trace positive
Infection intensity (POC-CCA, Tr-)	Infection categories from urine sample using POC-CCA test; trace negative
Village prev. (KK)	Village level prevalence of infection by KK
Village prev. $\geq 50\%$ (KK)	Village prevalence setting of schistosomiasis infection by KK.
Village prev. (POC-CCA, Tr-)	Village level prevalence of infection by POC-CCA trace negative.
Village prev. (POC-CCA, Tr+)	Village level prevalence of infection by POC-CCA trace positive; continuous
Village prev. $\geq 31\%$ (POC-CCA, Tr-)	Village prevalence setting of schistosomiasis infection by POC-CCA trace negative.
Village prev. (POC-CCA, Tr+)	Village prevalence for <i>S. mansoni</i> infection by POC-CCA trace positive.
Village prev. (≥ 1 EPG) & (POC-CCA, Tr-)	Village-level infection prevalence by agreement of KK and POC-CCA trace negative
Village prev. (≥ 1 EPG) & (POC-CCA, Tr+)	Village-level infection prevalence by agreement of KK and POC-CCA trace positive
Village prev. (KK, SAC only)	Village prevalence of infection (KK) among school-age children
Village prev. (POC-CCA, Tr-, SAC only)	Village prevalence of infection (POC-CCA trace negative) among school-age children
Village prev. (POC-CCA, Tr+, SAC only)	Village prevalence of infection (POC-CCA trace positive) among school-age children
Village prev. (KK, Adults only)	Village prevalence of infection (KK) in adults only
Village prev. (POC-CCA, Tr+, Adults only)	Village prevalence of infection (POC-CCA trace positive) among adults
Village prev. (POC-CCA, Tr-, Adults only)	Village prevalence of infection (POC-CCA trace negative) among adults
PHI $< 5\%$	Study participant lived in village with prevalence of heavy infection intensity (PHI; 400+ eggs per gram (EPG)) $< 5\%$ among school-aged children (SAC) and adults
PHI $< 1\%$	Study participants lived in village with PHI $< 1\%$ among SAC and adults
PHI $< 5\%$ (SAC only)	Study participants lived in village with PHI $< 5\%$ among SAC only
PHI $< 1\%$ (SAC only)	Study participants lived in village with PHI $< 1\%$ among SAC only
PHI $< 5\%$ (Adults only)	Study participants lived in village with PHI $< 5\%$ among adults only
PHI $< 1\%$ (Adults only)	Study participant lived in village with PHI $< 1\%$ among adults only

Table S2 Descriptive statistics of household and village-level variables

Variable	Overall N = 2834 n (%)	Non-PPF N = 2492 n (%)	PPF N = 342 n (%)
Home quality score, median (IQR)	6 (3-9)	6 (3-9)	3 (3-7-75)
Household uses safe drinking water source	2281 (80-49)	1990 (79-86)	291 (85-09)
Household has improved sanitation	2031 (71-67)	1788 (71-75)	243 (71-05)
Household has improved hygiene facilities	248 (8-75)	210 (8-43)	38 (11-11)
No. rooms in home, median (IQR)	2 (1-3)	2 (1-3)	2 (1-3)
2+ rooms in home	1972 (69-58)	1734 (69-58)	238 (69-59)
Own a home	2397 (84-58)	2093 (83-99)	304 (88-89)
Home has electricity	1089 (38-43)	964 (38-68)	125 (36-55)
Household purifies drinking water	636 (22-44)	543 (21-79)	93 (27-19)
Household has social status	280 (9-88)	236 (9-47)	44 (12-87)
Household head has social status	236 (8-33)	197 (7-91)	39 (11-40)
Household settlement (yrs), median (IQR)	15 (7-27)	15 (7-26)	20 (9-30)
Village prev. (KK), median (IQR)	41-25 (33-33-56-25)	41-25 (33-33-56-25)	45 (33-75-58-54)
Village prev. (KK)			
11-49	1739 (61-36)	1542 (61-88)	197 (57-60)
50-100	1095 (38-64)	950 (38-12)	145 (42-40)
Village prev. (POC-CCA, Tr-), median (IQR)	40-74 (22-50-53-75)	40-74 (26-25-53-75)	42-50 (20-61-25)
Village prev. (POC-CCA, Tr+), median (IQR)	68-75 (58-75-76-25)	68-75 (58-75-76-25)	66-25 (58-75-73-75)
Village prev. (POC-CCA, Tr-)			
≤30	869 (30-66)	756 (30-34)	113 (33-04)
31-74	1965 (69-34)	1736 (69-66)	229 (66-96)
Village prev. (POC-CCA, Tr+)			
≤30	79 (2-79)	75 (3-01)	4 (1-17)
31-74	1962 (69-23)	1709 (68-58)	253 (73-98)
≥75	793 (27-98)	708 (28-41)	85 (24-85)
Village prevalence (≥ 1 EPG) & (POC-CCA, Tr-), median (IQR)	26-25 (15-37-50)	26-25 (15-37-50)	27-50 (16-25-44-06)
Village prevalence (≥ 1 EPG) & (POC-CCA, Tr+), median (IQR)	35 (26-25-47-50)	35 (26-25-47-50)	36-25 (26-25-50)
PHI <5 (SAC only)	951 (33-56)	859 (34-47)	92 (26-90)
PHI <1 (SAC only)	432 (15-24)	378 (15-17)	54 (15-79)
Village prev. (KK, SAC only)			
≤10	80 (2-82)	67 (2-69)	13 (3-80)
11-49	1539 (54-30)	1378 (55-30)	161 (47-08)
≥50	1215 (42-87)	1047 (42-01)	168 (49-12)
Village prev. (POC-CCA, Tr-, SAC only)			
≤30	748 (26-39)	647 (25-96)	101 (29-53)
31-74	1848 (65-21)	1651 (66-25)	197 (57-60)
≥75-100	238 (8-40)	194 (7-78)	44 (12-87)
Village prev. ≥ 75 (POC-CCA Tr+, SAC only)	1464 (51-66)	1309 (52-53)	155 (45-32)
PHI <5 (Adults only)	753 (26-57)	697 (27-97)	56 (16-37)
PHI <1 (Adults only)	476 (16-80)	433 (17-38)	43 (12-57)
Village prev. (KK, Adults only)	900 (31-76)	784 (31-46)	116 (33-92)
Village prev. ≥ 75 (POC-CCA Tr-, Adult only)	1522 (53-71)	1321 (53-01)	201 (58-77)
Village prev. (POC-CCA, Tr+, Adults only)			
≤30	160 (5-65)	153 (6-14)	7 (2-05)
31-74	2043 (72-09)	1768 (70-95)	275 (80-41)
≥75-100	631 (22-27)	571 (22-91)	60 (17-54)
PHI <5	1193 (42-10)	1084 (43-50)	109 (31-87)
PHI <1	80 (2-82)	67 (2-69)	13 (3-80)
Water sites in village			
None	430 (15-17)	366 (14-69)	64 (18-71)
Beach	1029 (36-31)	931 (37-36)	98 (28-65)
Landing site only	1375 (48-52)	1195 (47-95)	180 (52-63)
Public latrine in village	1809 (63-83)	1559 (62-56)	250 (73-10)
District			
Mayuge	959 (33-84)	912 (36-60)	47 (13-74)
Buliisa	950 (33-52)	835 (33-51)	115 (33-63)
Pakwach	925 (32-64)	745 (29-90)	180 (52-63)

*No of households = 1442

Table S3 Descriptive statistics of villages-level infection indicators

Variable	N	n (%)
Village prev. (KK), median (IQR)	38	41·25 (33·44-55·94)
Village prev. (KK)	38	
11-49%		24·00 (63·16)
≥ 50%		14·00 (36·84)
Village prev. (POC-CCA, Tr-), median (IQR)	38	40·74 (26·88, 53·44)
Village prev. (POC-CCA, Tr+), median (IQR)	38	68·12 (59·06-75·94)
Village prev. ≥31 (POC-CCA, Tr-)	38	
≤ 30%		12·00 (31·58)
31-74%		26·00 (68·42)
Village prev. (POC-CCA, Tr+)	38	
≤ 30%		1·00 (2·63)
31-74%		26·00 (68·42)
≥ 75-100%		11·00 (28·95)
Village prevalence (≥ 1 EPG) & (POC-CCA, Tr-), median (IQR)	38	25·62 (16·88- 37·50)
Village prevalence (≥ 1 EPG) & (POC-CCA, Tr+), median (IQR)	38	35·62 (26·56-46·88)
PHI <5% (SAC only)	38	14·00 (36·84)
PHI <1% (SAC only)	38	7·00 (18·42)
Village prev. (KK, SAC only)	38	
≤ 10%		1·00 (2·63)
11-49%		21·00 (55·26)
≥ 50%		16·00 (42·11)
Village prev. (POC-CCA, Tr-, SAC only)	38	
≤ 30%		10·00 (26·32)
31-74%		25·00 (65·79)
≥ 75%		3·00 (7·89)
PHI <5% (Adults only)	38	10·00 (26·32)
PHI <1% (Adults only)	38	6·00 (15·79)
Village prev. (KK, Adults only)	38	12·00 (31·58)
Village prev. ≥ 75 (POC-CCA Tr+, SAC only)	38	20·00 (52·63)
PHI <5%	38	17·00 (44·74)
PHI <1%	38	1·00 (2·63)
Water sites in village	38	
None		6·00 (15·79)
Beach		14·00 (36·84)
Landing site only		18·00 (47·37)
Public latrine in village	38	24·00 (63·16)

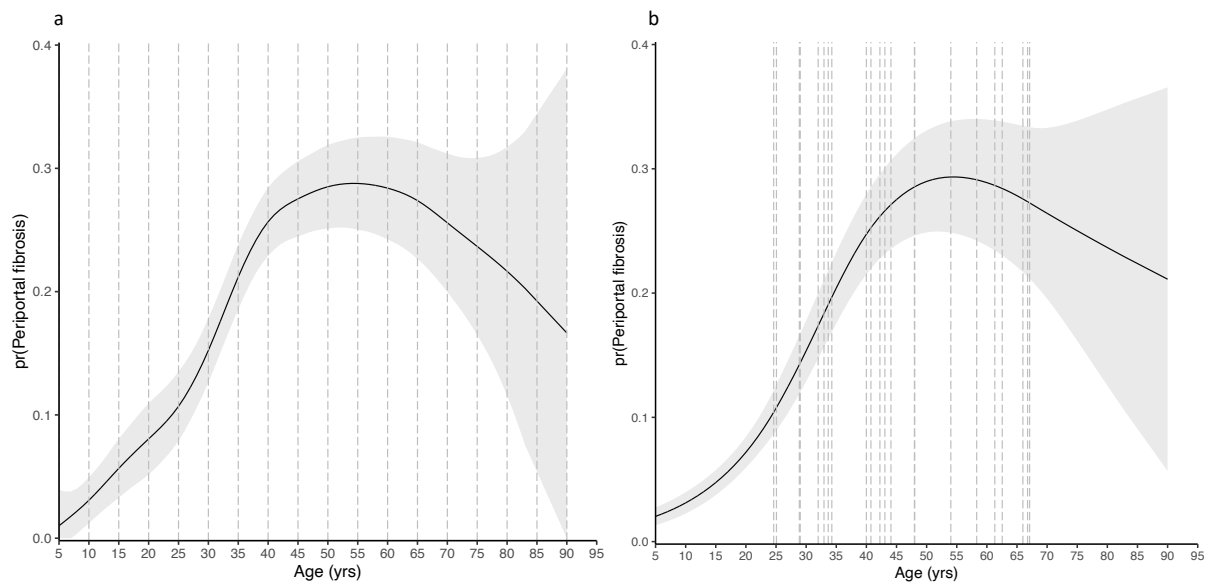


Figure S1 Generalised Additive Model (GAM) with fitted value (logit) of PPF over age. a. A GAM is shown with knots placed at every 5 years on the x-axis (age). **b.** A GAM is shown with 22 knots using the “freeknotsplines package in R”. It uses a random search algorithm and generalised cross validation (GCV) for selection of optimised knots numbers and their placement.⁷

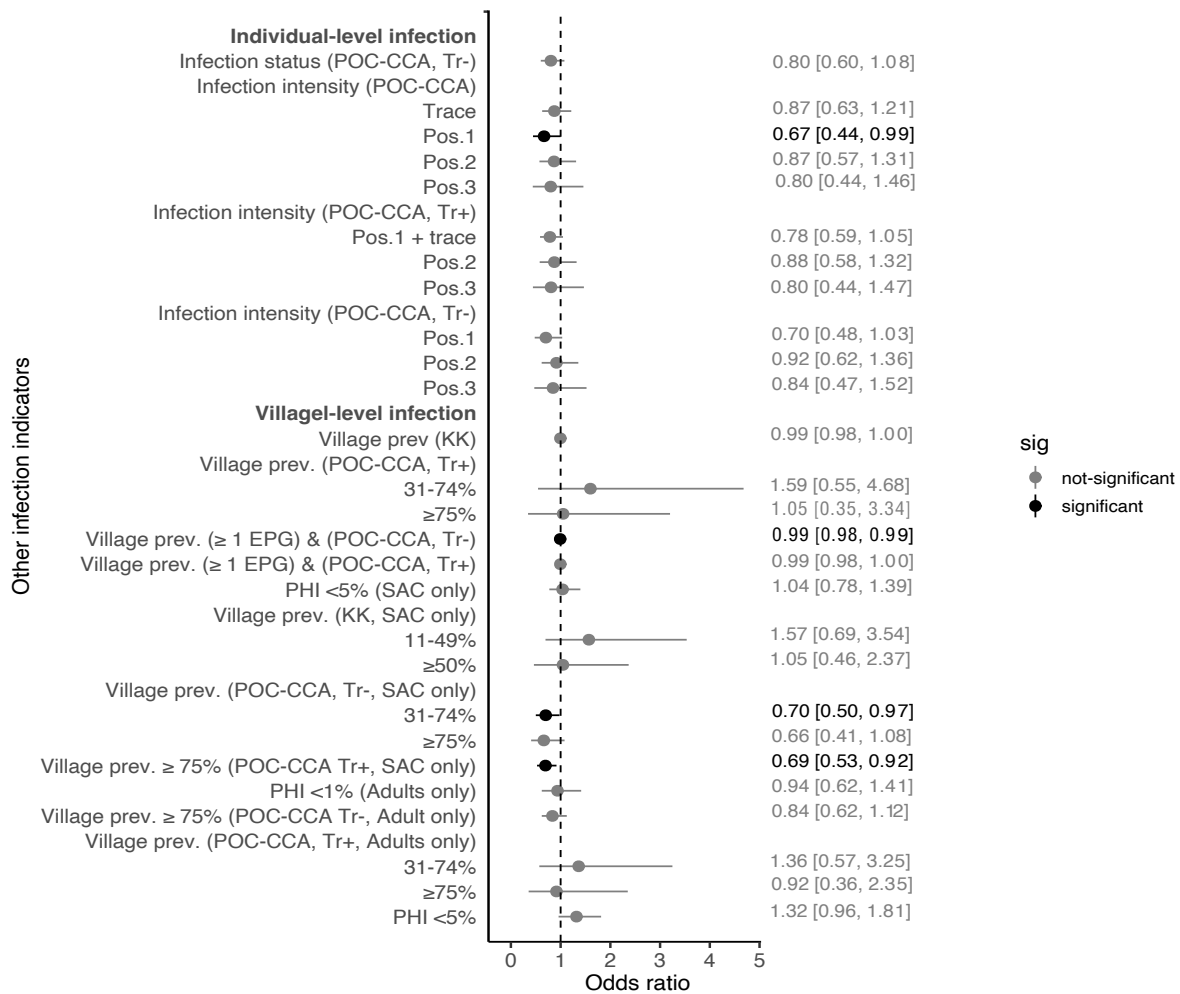


Figure S2 Fully adjusted models with alternative infection indicators. Alternative infection indicators, each representing a separate fully adjusted model, are shown. For each adjusted model, infection status (POC-CCA, Tr+) in Figure 6 was replaced by the other individual-level infection indicators each time while PHI <5% (Adults only) in Figure 6 was replaced by village-level infection indicators at the same time. All models were adjusted for age, gender, religion, tribe, all the medical history indicators, ultrasound-detected chronic hepatitis/early cirrhosis-like disease, malaria diagnosis, water contact, alcohol consumption and smoking status, occupation and all the household level variables in Figure 5. VIFs were between 1.1-2.3. The mean area under the ROC (10-fold CV) in each adjusted model ranged from 0.80-0.83.

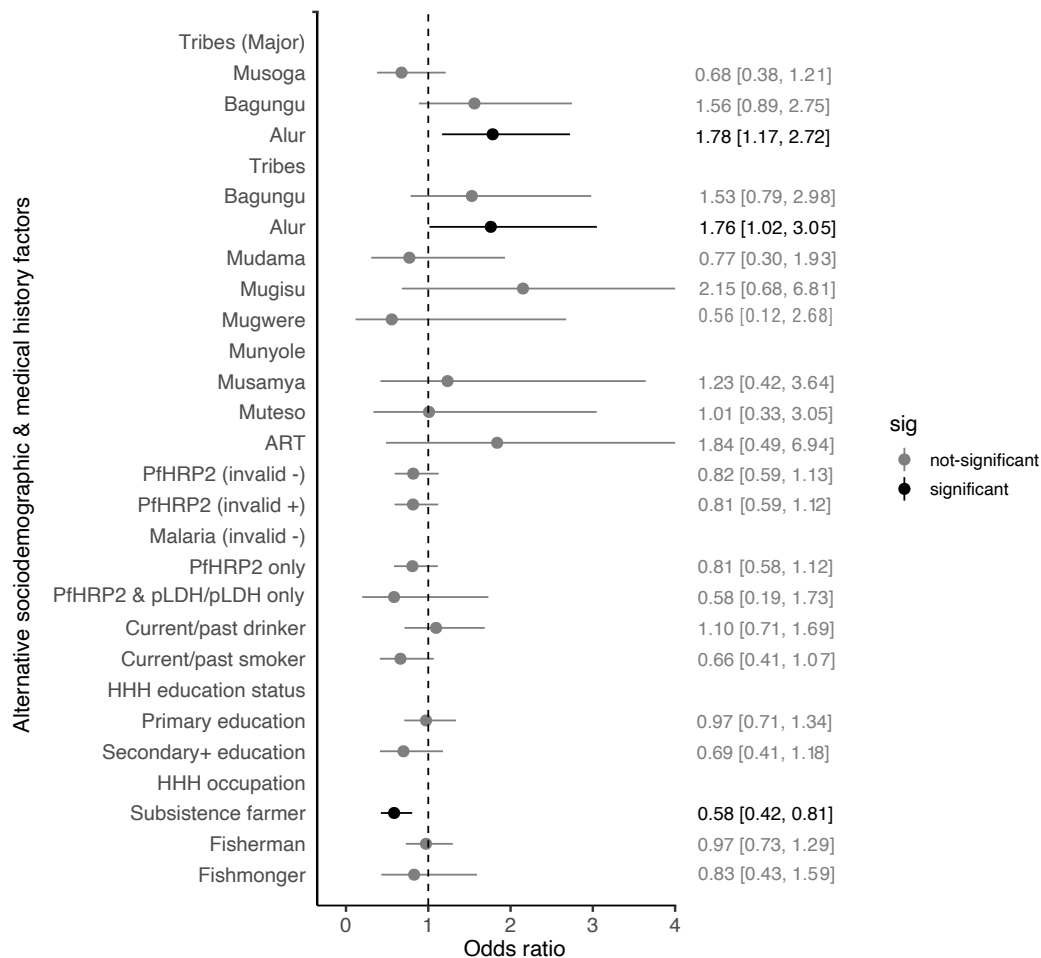


Figure S3 Fully adjusted models with alternative comorbidities and sociodemographic factors. Alternative definitions of sociodemographic and morbidity indicators, especially malaria, are shown where each represent a separate fully adjusted model. Each model was adjusted for age, gender, religion, tribe, all the medical history indicators, ultrasound-detected chronic hepatitis/early cirrhosis-like disease, malaria diagnosis, water contact, current drinker and current smokers, occupation and all the household level variables in Figure 5. VIFs were between 1.1-4.6. The mean area under the ROC (10-fold CV) in each adjusted model ranged from 0.81-0.84. PfHRP2 = Plasmodium falciparum antigen histidine rich protein 2. pLDH = parasite lactase dehydrogenase (generic to all malaria species). For pLDH only, there were only four cases.

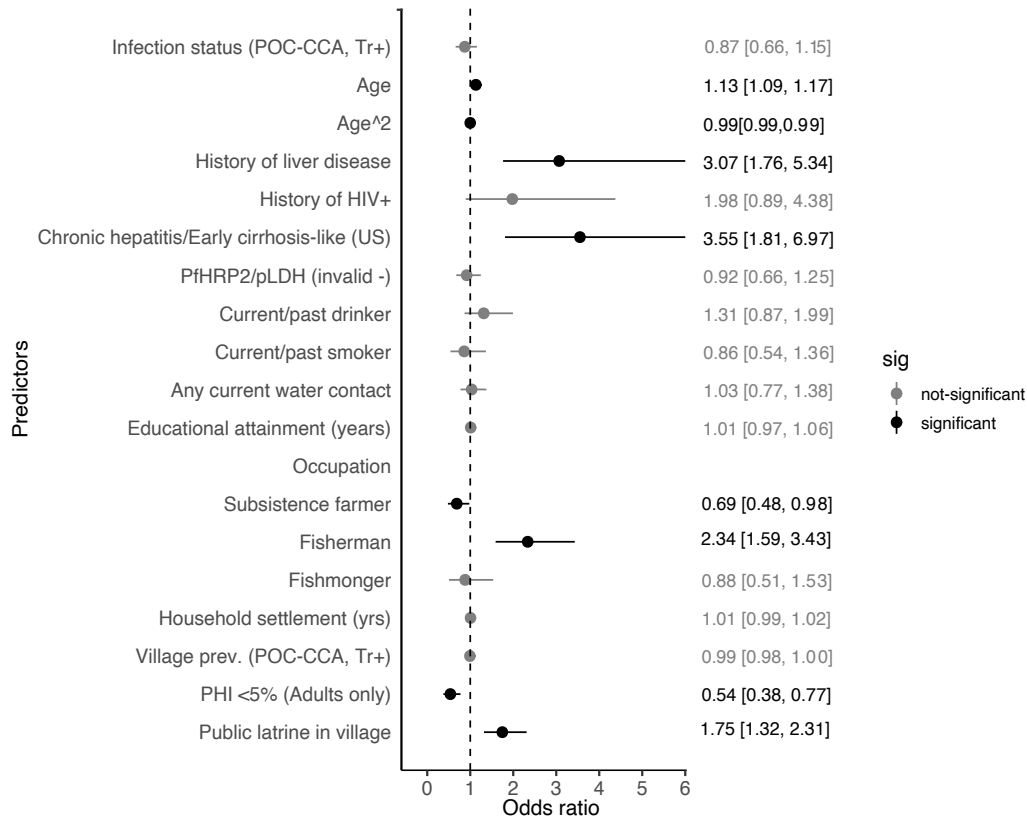


Figure S4 Subgroup analysis of participants from Pakwach District. Obs.= 925. Black represents significant relationships of p-value ≤ 0.05 . VIFs were < 10 . The mean area under the ROC (10-fold CV) was 0.81. PfHRP2 = Plasmodium falciparum antigen histidine rich protein 2. pLDH = parasite lactase dehydrogenase (generic to all malaria species). For pLDH only, there were only four cases. US = Ultrasound.

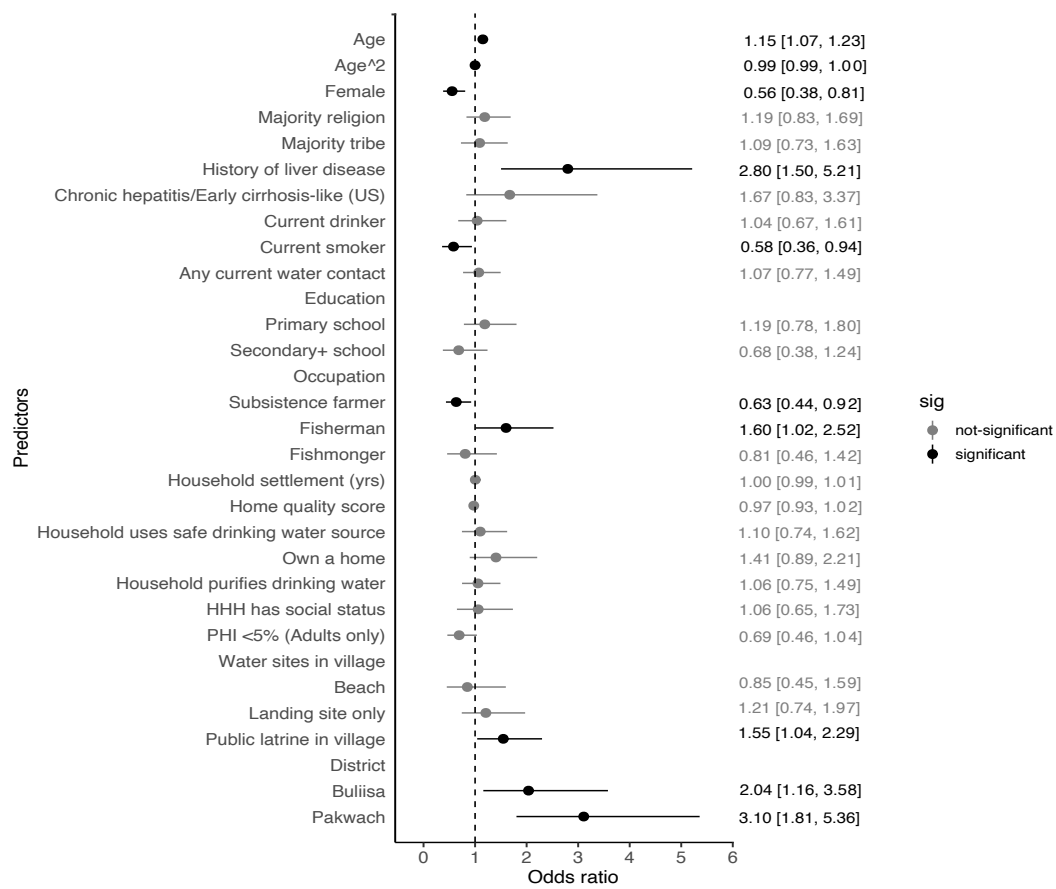


Figure S5 Subgroup analysis of adult participants. Obs. 1426. Black represents significant relationships of $p\text{-value} \leq 0.05$. VIFs were < 10 . The mean area under the ROC (10-fold CV) was 0.74. US = Ultrasound

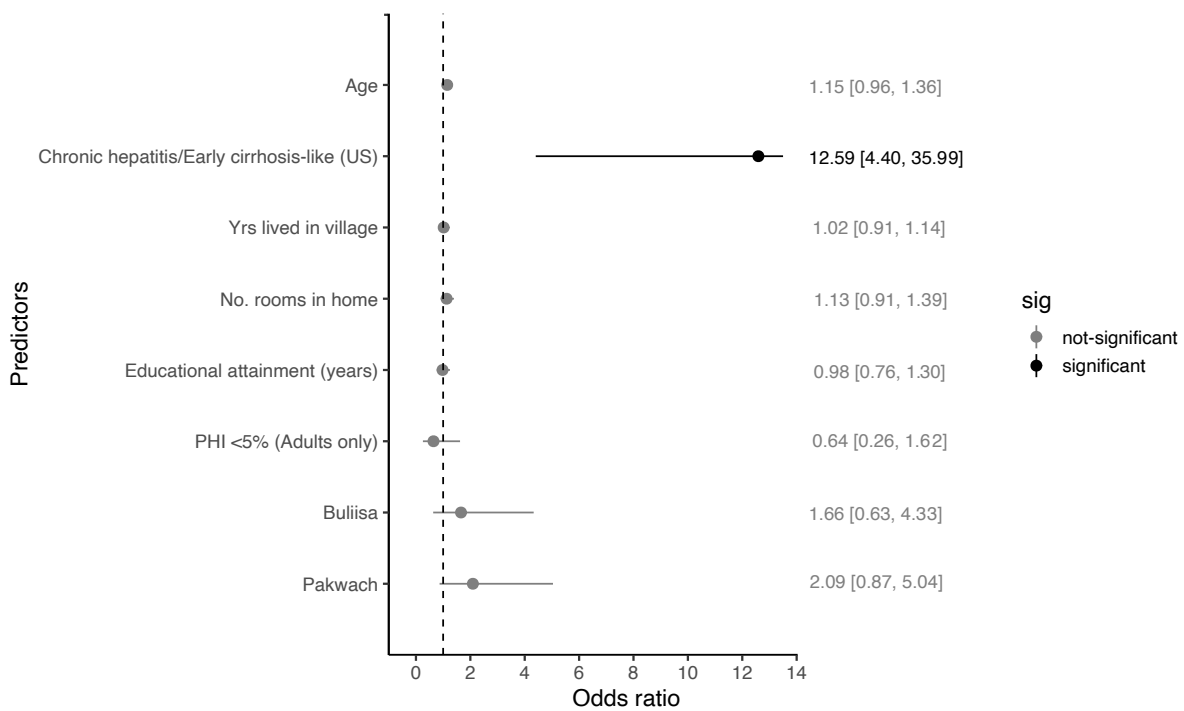


Figure S6 Subgroup analysis of children participants. Obs=1408. Black represents significant relationships of $p\text{-value} \leq 0.05$. VIFs were < 10 . The mean area under the ROC (10-fold CV) was 0.70. US = Ultrasound

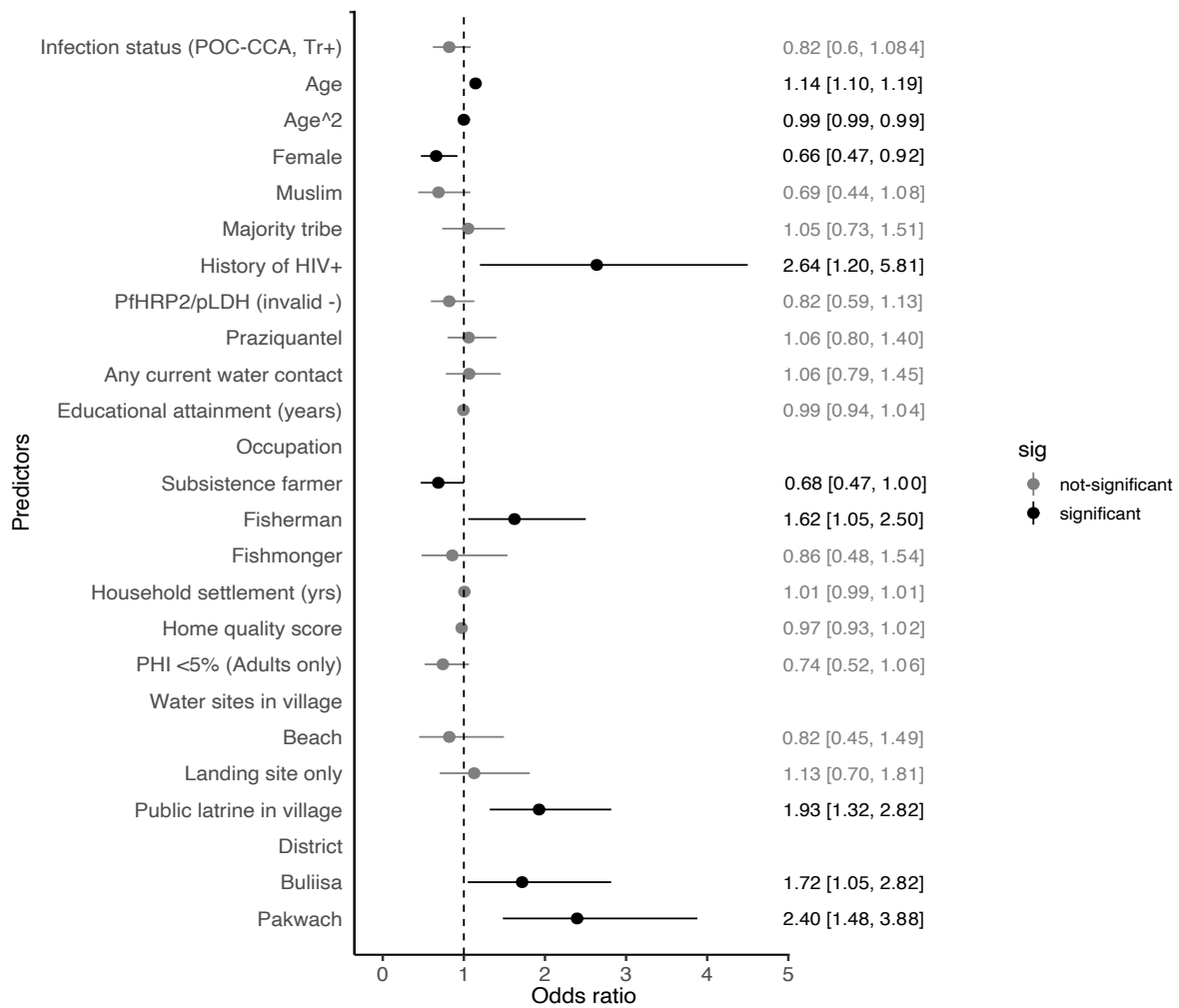


Figure S7 Sensitivity analysis of XY liver patterns. Obs.=2834. Study participants with liver pattern XY (the presence of a comorbidity including ultrasound-detected chronic hepatitis/early cirrhosis-like disease, fatty-like liver or a history of any liver diseases were recoded as zero (no PPF). Black represents significant relationships of $p\text{-value} \leq 0.05$. VIFs were < 10 . The mean area under the ROC (10-fold CV) was 0.81. PfHRP2 = *Plasmodium falciparum* antigen histidine rich protein 2. pLDH = parasite lactase dehydrogenase (generic to all malaria species). For pLDH only, there were only four cases.

References

1. World Health Survey (WHS). WHO Multi-Country Studies Data Archive: an Online Microdata Catalog. 2023. Available at: <https://apps.who.int/healthinfo/systems/surveydata/index.php/catalog/whs>. Accessed 11 July 2023. Accessed 28 June, 2023
2. World Health Organization. Noncommunicable disease surveillance, monitoring and reporting.: STEPwise approach to NCD risk factor surveillance (STEPS). Available at: <https://www.who.int/teams/noncommunicable-diseases/surveillance/systems-tools/steps>. Accessed 14 July 2023.
3. Tchelepi H, Ralls PW, Radin R, Grant E. Sonography of diffuse liver disease. *J Ultrasound Med* 2002; **21**:1023–1032.
4. Gaiani S, Gramantieri L, Venturoli N, et al. What is the criterion for differentiating chronic hepatitis from compensated cirrhosis? a prospective study comparing ultrasonography and percutaneous liver biopsy. *J. Hepatol.* 1997; **27**:979–985.
5. Ferraioli G, Monteiro LBS. Ultrasound-based techniques for the diagnosis of liver steatosis. *World J Gastroenterol* 2019; **25**:6053–6062.
6. Chami GF, Kontoleon AA, Bulte E, et al. Profiling Nonrecipients of Mass Drug Administration for Schistosomiasis and Hookworm Infections: A Comprehensive Analysis of Praziquantel and Albendazole Coverage in Community-Directed Treatment in Uganda. *Clin Infect Dis* 2016; **62**:200–7.
7. Spiriti S, Eubank R, Smith PW, et al. Knot selection for least-squares and penalized splines. *JSCS* 2013; **83**:1020–36. doi:10.1080/00949655.2011.647317