Variable	Type	Label	Variable definition/code	description
school_id	int			101 distinct 4 digit codes
child_id	int	Child ID		13527 distinct 7 digit codes. 1st 4 digits are school code and last 3 are individual child numbers
indicator	category		1=single child observation	restricts the data to individual observations of all 13527 study children
study_grp	categorical	Intervention group	1=malaria and literacy intervention 2=malaria intervention 3=literacy intervention 4=control	
MAL_grp	binary	Malaria intervention	0=no 1=yes	0=not in malaria intervention group 1=in malaria intervention group
LIT_grp	binary	Literacy intervention	0=no 1=yes	0=not in literacy intervention group 1=in literacy intervention group
class_m	binary	Class	1=class 1 5=class 5	classes at enrolment into the study
refused	binary	parent refused consent	0=no 1=yes	
consent	binary	child has consent	0=no 1=yes	parent provided written informed consent
study_child	binary	Study child (>1 educ task + anaemia (all schools) + slide (malaria int sch)	0=no 1=yes	enrolled into the trial initially
withdrawn	binary	Withdrawn during study	0=no 1=yes	
final_withdraw_date_formatted	stata date	Withdrawal date (formatted) - FINAL		
dead	binary	Died during follow-up	0=no 1=yes	
date_death_final_formatted	stata date	Date of death (approx) - formatted		
withdrawn_dead	binary	Study child: dead or withdrawn during follow-up	0=no 1=yes	
bl_edassess	binary	BL Ed assessed (C1+C5)	0=<2 BL education tasks 1=>1 BL education task	child required to have one or more education assessment tasks at baseline to be defined as a study child

Variable	Type	Label	Variable definition/code	description
withdrawn_bl	binary	child withdrawn at baseline data collection	0=no 1=yes	withdrew during the baseline education or health analyses and thus was excluded from the baseline analysis and from subsequent data collection and analyses.
bl_hlthassess	binary	Health assessed i.e. anaemia (& fal if MAL INT) at baseline (only for those not	0=no 1=yes	child required to have both Hb and microscopy result to be defined as health assessed at baseline.
FU1_edassess	binary	Study child + not withdrawn/dead FU1 educ + FU1 Ed assessed (C1+C5)	0=0 FU1 education tasks 1=>0 FU1 education task	has to have one of the three educational assessments - literacy, numeracy or attention to be defined as being educationally assessed at FU1
FU1_hlthassess	binary	Study child + not withdrawn/dead FU1 health + anaemia or falciparum	0=no 1=yes	has to have either anaemia or malaria outcome to be defined as being health assessed
FU2_edassess	binary	FU2 educ (spell or attention or numeracy) available (accounting for withdrawals)	0=no 1=yes	has to have one of the three educational assessments - literacy, numeracy or attention to be defined as being educationally assessed at FU2
FU2_healthassess	binary	Study child + not withdrawn/dead at FU2 health + anaemia or falciparum (i.e. at	0=no 1=yes	has to have either anaemia or malaria outcome to be defined as being health assessed
age_child	continuous	Age at BL (TO BE USED FOR ANALYSIS)		child reported age
age_cat2	categorical	Age cat (child-reported) - used in BL paper	0= 5-9 1= 10-12 2= 13-20	in years
age_child_fu1	continuous	age at FU1		age at baseline plus one year - used to calculate anaemia at FU1
age_child_fu2	continuous	age at FU2		age at baseline plus two years - used to calculate anaemia at FU2
sex	binary	sex	1=male 2=female	
BL_numeracy_younger	continuous	Class 1 Arithmetic: Addition (score: 0-30)	0-30	Baseline numeracy outcome in class 1
BL_numeracy_older	continuous	Class 5 Numeracy: written numeracy (score: 0-38)	0-38	Baseline numeracy outcome in class 5
BL_attention_younger	continuous	Class 1 Attention: Pencil tap	0-20	Baseline sustained attention outcome class 1

Variable	Type	Label	Variable definition/code	description
BL_attention_older	continuous	Class 5 Attention: Double digit code transmision (score: 0-20)	0-20	Baseline sustained attention outcome class 5
BL_spelling_younger	continuous	Class 1 Literacy: Spelling(score: 0-20)	0-20	Baseline literacy outcome in class 1
BL_spelling_older	continuous	Class 5 Literacy: Spelling - ftr pts + total worlds correct (score: 0-78)	0-78	Baseline literacy outcome in class 5
BL_hb	continuous	baseline haemoglobin concentration		
BL_malaria	binary	baseline RDT result	0=RDT negative 1=RDT positive	
BL_height	continuous	baseline height in cm		
BL_weight	continuous	baseline weight in kg		
BL_temp	continuous	baseline auxiliary temperature in celcius		
BL_falciparum	binary	P.falciparum (positive by both microscopy readers or resolved by 3rd)	0=negative 1=positive	
BL_anaemia	binary	Anaemia (based on child-reported age at baseline)	0=not anaemic 1=anaemic	age and sex correct based on WHO thresholds: haemoglobin concentration <110g/l in children under 5 years; <115g/l in children 5 to 11 years; <120g/l in females 12 years and over and males 12 to 15 years old; and <130g/l in males over 15 years
BL_stunted	binary	child is stunted?	0=no 1=yes	stunting defined as >2SD from the reference median in ANTHROPLUS
BL_underweight	binary	child is underweight?	0=no 1=yes	underweight defined as >2SD from the reference median in ANTHROPLUS
BL_thin	binary	child is thin?	0=no 1=yes	thinness defined as >2SD from the reference median in ANTHROPLUS
BL_anaemia_cat	categorical	Anaemia category	0=Severe (< 70 g/l) 1=Moderate (70-89 g/l) 2=Mild (90-109 g/l) 3= None (>=110 g/l)	based on crude Hb values

Variable	Type	Label	Variable definition/code	description
BL_anaemia_mild_mod_severe	binary	Mild-moderate-severe anaemia (<110 g/l)	0=not anaemic 1=anaemic	based on crude cut off of <110g/l
BL_ses	categorical	Baseline SES (from PCA - as used in BL paper)	1=poorest 2=poor 3=Median poor 4=less poor 5=least poor	based on PCA
net_child	binary	Child sleeps under a net	0=no 1=yes	
lastnightnet_child	binary	Child sleeps under a net last night	0=no 1=yes	only asked to those children who sleep under a net
bicycle	binary	Household owns a bicycle	0=no 1=yes	
motorcycle	binary	Household owns a motorcycle	0=no 1=yes	
radio	binary	Household owns a radio	0=no 1=yes	
television	binary	Household owns a TV	0=no 1=yes	
mobilephone	binary	Household owns a mobilephone	0=no 1=yes	
electricity	binary	Household has electricity	0=no 1=yes	
pitlatrine	binary	Household has a pit latrine	0=no 1=yes	
sfp	binary	School feeding programme	0=no 1=yes	school-level variable
deworm	binary	School deowrming done in last 12 months	0=no 1=yes	school-level variable
mal_control	binary	Malaria control programme	0=no 1=yes	school-level variable
division	categorical	Division	1=Diani 2=Lunga Lunga 3=Msambweni 4=Kubo	district division names

Variable	Type	Label	Variable definition/code	description
cluster	string	School-cluster (unit randomisation to literacy intervention)	names of the 26 clusters	
mal_pot	continuous	stratification for malaria group randomisation	0-1	
kcpemean_mal	continuous	School mean exam score (2008)	166-320	school-level variable
cluster_size2	digit	cluster size	3-6	number of schools in each TAC cluster
cluster_KCPE_mean	continuous	Cluster KCPE (mean of all schools in cluster)	26 x mean scores	
schlevel_comp	categorical	Education of household head	0=no schooling 1=primary 2=secondary 3=college/degree	
crowding	categorical	number of people in house	1=1-5 2=6-7 3=8-9 4=10-32	
brick_walls	binary	house has brick/cement walls?	0=no 1=yes	
cement_floor	binary	house has cement floors?	0=no 1=yes	
iron_roof	binary	house has iron/tile roof?	0=no 1=yes	
cov_water	binary	water supply for house?	0=uncovered 1=covered	
visit	binary	Follow-up visit	1=Follow-up 1 2=Follow-up 2	long format
attention_younger	continuous	Class 1 Attention FOR ANALYSIS: single digit code transmission	0-20	for both visit 1 and 2 in long format
attention_older	continuous	Class 5 Attention FOR ANALYSIS: double digit code transmission	0-20	for both visit 1 and 2 in long format
spelling_younger	continuous	Class 1 Spelling FOR ANALYSIS (score: 0-20)	0-20	for both visit 1 and 2 in long format

Variable	Type	Label	Variable definition/code	description
spelling_older	continuous	Class 5 Spelling FOR ANALYSIS (score: 0-78)	0-78	for both visit 1 and 2 in long format
numeracy_younger	continuous	Class 1 Numeracy FOR ANALYSIS: Arithmetic -addition(0-30 at FU1), written (0-38 at FU2)	0-30 & 0-38	for both visit 1 and 2 in long format
numeracy_older	continuous	Class 5 Numeracy FOR ANALYSIS: written (0-38)	0-38	for both visit 1 and 2 in long format
FU_health_status	categorical	Health follow-up (0= neither; 1= FU1 only; 2 = FU2 only; 3 = both)	0=Neither 1=FU1 only 2=FU2 only 3=FU1 and FU2	
FU_educ_status	categorical	Educ follow-up - spelling/attention/numeracy (0= neither; 1= FU1 only; 2 = FU2 only)	0=Neither 1=FU1 only 2=FU2 only 3=FU1 and FU2	
study_child_FU1_educ_analysis	binary	Study child with consent + not withdrawn/dead FU1 educ	0=no 1=yes	
study_child_FU1_health_analysis	binary	Study child with consent + not withdrawn/dead FU1 health	0=no 1=yes	
study_child_FU2_analysis	binary	Study child with consent + not withdrawn/dead FU2 educ	0=no 1=yes	
FU_edassess	binary	FU educ (spell, numer or attention), visit-specific (accounts withdraw/dead)	0=no 1=yes	variable is specific for each follow-up
FU_hlthassess	binary	FU health (Pf or anaemia), visit-specific (accounts withdrawn/dead)	0=no 1=yes	variable is specific for each follow-up
fu1_educ_withdrawn	binary	label value lastnightnet_child lastnightnet_child_lab	0=no 1=yes	
fu1_health_withdrawn	binary	Withdrawn at FU1 health analysis	0=no 1=yes	
fu2_educ_withdrawn	binary	Withdrawn at FU2 educ analysis	0=no 1=yes	
fu2_health_withdrawn	binary	Withdrawn at FU2 health analysis	0=no 1=yes	
dead_9mth_educ	binary	Death status at 9mth educ i.e. FU1 (for eventual deaths)	0=no 1=yes	

Variable	Type	Label	Variable definition/code	description
dead_12mth_health	binary	Death status at 12mth health i.e. FU1 (for eventual deaths)	0=no 1=yes	
dead_24mth_educ	binary	Death status at 24mth educ i.e. FU2 (for eventual deaths)	0=no 1=yes	
dead_24mth_health	binary	Death status at 24mth health i.e. FU2 (for eventual deaths)	0=no 1=yes	
fal_categories_INFERRED_3grps	categorical	BASELINE falciparum prev categories (using BL and FU1 data for INT and control)	1=<5% 2=5-19.9% 3=>=20%	
falciparum	binary	P.falciparum (positive by both microscopy readers or resolved by 3rd)	0=negative 1=positive	for both visit 1 and 2 in long format
anaemia	binary	Age-specific anaemia (using visit-specific age, based on child-reported BL age)	0=not anaemic 1=anaemic	for both visit 1 and 2 in long format
anaemia_cat	categorical	Anaemia category	0=Severe (< 70 g/l) 1=Moderate (70-89 g/l) 2=Mild (90-109 g/l) 3= None (>=110 g/l)	for both visit 1 and 2 in long format
hb	continuous	haemoglobin concentration g/l		for both visit 1 and 2 in long format
height	continuous	height in cm		for both visit 1 and 2 in long format
weight	continuous	weight in kg		for both visit 1 and 2 in long format
temp	continuous	auxiliary temperature in celcius		for both visit 1 and 2 in long format
term_of_transfer	categorical	First visit number of transfer	terms 1 to 7	term at which child was noticed transferred
FU1_ed_include_TRANSFER	binary	FU1 ed: study child not withdrawn/dead/transferred	0=no 1=yes	if you want to run analysis excluding transfers
FU1_health_include_TRANSFER	binary	FU1 health: study child not withdrawn/dead/transferred	0=no 1=yes	if you want to run analysis excluding transfers
FU2_ed_include_TRANSFER	binary	FU2 ed: study child not withdrawn/dead/transferred	0=no 1=yes	if you want to run analysis excluding transfers
FU2_health_include_TRANSFER	binary	FU2 health: study child not withdrawn/dead/transferred	0=no 1=yes	if you want to run analysis excluding transfers